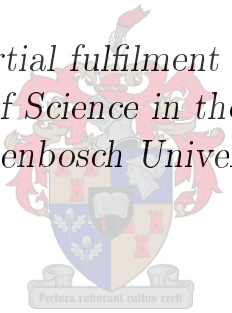
Structural Emergence in Mutualistic Networks

by

Chinenye Assumpta Nnakenyi

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in the Faculty of Science at Stellenbosch University



Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Supervisor: Prof. C. Hui

December 2016

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: .....
C.A. Nnakenyi

Date: .....December 2016.....

Abstract

Structural Emergence in Mutualistic Networks

C.A. Nnakenyi

Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Thesis: M.Sc

September 2016

Mutualistic interactions are vital in sustaining species, maintaining the functions and services of the ecosystem. Network structures such as nestedness and modularity have being reported to emerge as a result of the mutualistic interactions. Although these structures have being found to have effect in the stability of mutualistic communities, mechanisms that lead to the emergence of these structures are not fully understood. From the observed pollination data of 10 Galápagos Islands, we use a modified Lotka-Voltera model of mutualism that incorporates species competition, functional responses and adaptive rewiring (Adaptive Interaction Switching [AIS] model) to predict the observed network structures. From the AIS model, almost 40% variation of the observed nestedness and more than 50% variation of the observed modularity was explained. Furthermore, using a Generalized Linear Model (GLM), the effect of environmental variables such as geographic factors (island area, isolation, age and maximum elevation) and anthropogenic factors (sampling effort and human population size) were considered together with the AIS model. The GLM can account for more than 78% variation of the observed nestedness and more than 85% of the observed modularity, with island area, isolation, sampling effort and human population size the most important variables, contributing significantly to the observed network structures. Therefore, pollination networks on Galápagos Islands are structured. The AIS model implemented can explain an appreciable level of network structure. Together with environmen-

tal variables, the results echo the importance of island area and isolation (as of island biogeography), human disturbance, sampling effort, as well as the adaptive rewiring (ecological fitting), as a candidate model for mutualistic network emergence.

Uittreksel

Strukturele Opkoms in Mutualistiese Netwerke

("Structural Emergence in Mutualistic Networks")

C.A. Nnakenyi

Departement Wiskundige Wetenskappe,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: M.Sc

September 2016

Mutualistiese interaksie is noodsaaklik vir die handhawing van spesies en die onderhoud van funksies en dienste van die ekosisteem. Netwerkstrukture soos genestheid en modulariteit is al berig om as 'n resultaat van sulke mutualistiese interaksies na vore te kom. Alhoewel hierdie strukture bevind is om 'n effek te hê op die stabiliteit van mutualistiese gemeenskappe, word meganismes wat lei tot die opkoms van hierdie strukture nog nie goed verstaan nie. Deur 10 bestuiwingsdata van die Galápagos-eilande the gebruik, het ons 'n aangepaste Lotka-Volterra model van mutualisme gebruik wat kompetisie tussen spesies, funksionele reaksies en aanpasbare herbedrading [rewiring] (Adaptive Interactive Switching [AIS] model) inkorporeer om die waargenome netwerkstrukture te voorspel. Uit die AIS model is byna 40% variasie van die waargenome genstheid en meer as 50% variasie van die waargenome modulariteit verduidelik. Verder, deur 'n Veralgemeende lineêre model (of GLM vir Generalized Linear Model) te gebruik, is die effek van geografiese faktore (eilandoppervlakte, isolasie, ouderdom en maksimum hoogte bo seespieël) en antropogeniese faktore (poging tot steekproefneming en menslike bevolkingsgrootte) saam met die AIS model in ag geneem. Die GLM kan tot 78% variasie van die waargenome genestheid en meer as 85% van die waargenome modulariteit verduidelik, met eliandoppervlakte, isolasie, poging tot steekproefneming en menslike bevolkingsgrootte die belangrikste veranderlikes wat aansienlik bydrae tot die

waargenome netwerkstrukture. Daarom is netwerke op die Galápagos-eilande gestruktureerd. Die AIS model wat ons geïmplememteer het, kan 'n aansienlike vlak van netwerkstrukture verduidelik. Saam met omgewingsveranderlikes weerspieël die bevindings die belang van eilandoppervlakte en isolasie (soos van eiland biogeografie), menslike versteuring, poging tot steekproefneming sowel as aanpasbare herbedrading (ekologiese passing) as 'n kandidaatmodel vir mutualistiese netwerkopkoms.

Acknowledgements

I thank God for His mercy, grace and strength to be able to start and complete my thesis, may His name be praised forever (Amen).

My sincere appreciation goes to my supervisor Prof. Cang Hui (SARCHI Chair) for his guidance, mentor and supervision of my thesis. I apprectiate Dr. Henintsoa Onivola Minoarivelo for her assistance, effort and encouragements all through my study period. I equally thank Savannah Nuwagaba for her assistance. Thanks to Vanessa Du Plessis and Dr. Dirk Basson for translating my abstract to Afrikaans. May God bless and reward you all.

I appreciate the effort of Prof. Anna Traveset and other collaborators for providing the mutualistic network data I used for analysis in this thesis.

I thank my family and friends for their endless support and fervent prayers for my success.

Lastly, I express my gratitude to Stellenbosch University, African Institute for Mathematical Sciences (AIMS) and my supervisor, Prof. Cang Hui, for the financial support they provided me with, helping my thesis to be a success.

Dedications

To my Father, Bernard, my late Mother, Paulina and my siblings, Lucy, John, Mary, Augustina, Peace and Emmanuel

Contents

Declaration i

Abstract ii

Uittreksel iv

Acknowledgements vi

Dedications vii

Contents viii

List of Figures xi

List of Tables xiv

1 Introduction 1
- 1.1 Background to the research problem . . . 1
- 1.2 Problem statement . . . 3
- 1.3 Research question . . . 3
- 1.4 Aims and objectives . . . 4
- 1.5 Thesis overview . . . 4

2 Literature Review 5
- 2.1 Ecological Interactions . . . 5
- 2.2 Complex Networks . . . 6
  - 2.2.1 Graph . . . 8
    - 2.2.1.1 Types of graph . . . 9
  - 2.2.2 Properties of complex networks . . . 9
    - 2.2.2.1 Small worlds . . . 9
    - 2.2.2.2 Clustering . . . 10
    - 2.2.2.3 Degree distribution . . . 10
    - 2.2.2.4 Network resilience . . . 10
  - 2.2.3 Network models . . . 10
    - 2.2.3.1 Random network models . . . 11

2.2.3.2 Small world models . . . 11
2.2.3.3 Scale free models . . . 11
2.3 Structures of Complex Networks . . . 12
2.3.1 Structures of Mutualistic Networks . . . 12
2.3.1.1 Community size . . . 13
2.3.1.2 Connectance . . . 13
2.3.1.3 Strength and asymmetry . . . 14
2.3.1.4 Degree distribution . . . 15
2.3.1.5 Nestedness . . . 15
2.3.1.6 Modularity . . . 16
2.4 Dynamic Network Models . . . 17
2.4.1 Functional Responses . . . 18
2.4.2 Adaptive Interaction Switch (AIS) models . . . 20
2.5 Environmental Variables . . . 21
2.6 Measures of Stability . . . 23
2.6.1 Asymptotic Stability . . . 23
2.6.2 Resilience . . . 23
2.6.3 Persistence . . . 24
2.6.4 Robustness . . . 24

3 Adaptive Interaction Switch Models and Study system 25
3.1 Galápagos Islands/Study data . . . 25
3.2 Model Development . . . 27
3.2.1 Dynamical description . . . 27
3.2.2 Adaptive Rewiring Algorithms . . . 29
3.2.3 Numerical Simulations and Parametrizations . . . 30
3.2.4 Null models . . . 31
3.2.5 Results . . . 32
3.2.5.1 Nestedness . . . 32
3.2.5.1.1 Different initial values convergence: . . . 32
3.2.5.1.2 Nestedness predictions: . . . 36
3.2.5.2 Modularity . . . 39
3.2.5.2.1 Different initial values convergence: . . . 39
3.2.5.2.2 Modularity predictions: . . . 40
3.2.5.3 Persistence . . . 43
3.2.6 Discussion . . . 44

4 Statistical Analysis of the Environmental Variables 46
4.1 Data . . . 46
4.2 Method of Analysis . . . 47
4.3 Results . . . 47
4.3.1 Nestedness . . . 47
4.3.1.1 Variance Inflation Factor (VIF) . . . 47
4.3.1.2 Multiple regression . . . 48

4.3.2 Modularity 51
4.3.2.1 Variance Inflation Factor (VIF) 51
4.3.2.2 Multiple regression 52
4.4 Discussion 55

5 Conclusions **57**
5.1 Summary 57
5.2 Limitation and further work 58

Appendices **59**

A R codes **60**

List of References **69**

List of Figures

2.1 Mutualistic interactions between plants and animals. (a)-(e) are pollination and (f) is seed dispersal (Severns (2016); Cole (2016); Tribune (2016); Batra (2016); Tutorvista (2016); Moore (2001)). . . 7
2.2 Ecological interactions: (a) and (b) are competition; (c) and (d) are antagonistic interactions (Mohammed (2016); Manaster (2016); Bar-Yam (2016*a*); Bar-Yam (2016*b*)). 8
2.3 Graph types 9
2.4 Genovesa island network 14
2.5 Network structures representing nestedness and modularity, where each box represents an interaction between plants and pollinators. (a) is a perfectly nested network, having interaction A which is between a generalist plant and a generalist pollinator, and interaction B, between a specialist plant and a generalist pollinator. (b) is a compartmentalised network. 16
2.6 Types of functional responses. 19
2.7 Island biogeography theory showing the immigration and extinction rates with species richness. E_1 is the equilibrium point for species richness in far and small islands, E_2 is the equilibrium point for species richness in close and small islands, E_3 is the equilibrium point for species richness in far and large islands and E_4 is the equilibrium point for species richness in close and large islands. . . 22

3.1 Galapagos Islands. 26
3.2 Nestedness and Modularity dynamics of the four rewiring algorithms (Adaptive Interaction Switch models). Model simulations are done using Española island data. Red broken line is the observed value of each structure. Elimination models are IE (type I functional response) and IIE (type II functional response); optimization models are IO (type I functional response) and IIO (type II functional response). 33

3.3 Nestedness and Modularity dynamics of the four rewiring algorithms (Adaptive Interaction Switch models). Model simulations are done using Santacruz island data. Red broken line is the observed value of each structure. Elimination models are IE (type I functional response) and IIE (type II functional response); optimization models are IO (type I functional response) and IIO (type II functional response). 34
3.4 Nestedness dynamics of Española Island using the elimination algorithm models (IE and IIE). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of nestedness. Model prediction fitted the observed level. 34
3.5 Nestedness dynamics of Santacruz Island using elimination algorithm models (IE and IIE). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of nestedness. Model prediction underestimated the observed level. . . 35
3.6 Nestedness dynamics of Española Island using optimization algorithm models (IO and IIO). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of nestedness. Model prediction overestimated the observed level. . . . 35
3.7 Nestedness predictions comparison between each of our four Adaptive Interaction Switch (AIS) models and the two null models against the observed values. Each dot represents an island network. Elimination models are IE (linear functional response) and IIE (nonlinear functional response); optimization models are IO (linear functional response) and IIO (nonlinear functional response); null models are Null4 and Null5. The broken line is for $y = x$. 38
3.8 Modularity dynamics of Santa Cruz Island using Elimination algorithm models (IE and IIE). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of modularity. Model prediction fitted the observed level. 39
3.9 Modularity dynamics of Santa Cruz Island using Optimization algorithm models (IO and IIO). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of modularity. Model prediction underestimated the observed level. . . 40

3.10 Modularity predictions comparison between each of our four Adaptive Interaction Switch (AIS) models and the two null models against the observed values. Each dot represents an island network. Elimination models are IE (linear functional response) and IIE (nonlinear functional response); optimization models are IO (linear functional response) and IIO (nonlinear functional response); null models are Null4 and Null5. The broken line is for $y = x$. 42
3.11 Log(x) plot of species population densities of the Adaptive Interaction Switch models using Española Island data. The left column figures are for plant species while the right column figures are for animal species. The first and third row figures are for optimization models IO and IIO respectively; the second and fourth row figures are for elimination models IE and IIE respectively. 43
3.12 Correlation of the observed connectance with observed nestedness and modularity, r is the correlation coefficient. 45

4.1 Nestedness variance partitioning showing the independent contribution of each of the explanatory variables to the total variance explained for different model combinations. The environmental variables are SE - sampling effort, ND - nearest distance (isolation), AR - area and HPS - human population size (log transformed [log(x+1)]). The band strips shows the 95% confidence interval of the randomly divided variance explained. Bars above this band are important and asterisked bars are significant ($p < 0.1$). Elimination models are IE (linear functional response) and IIE (nonlinear functional response); optimization models are IO (linear functional response) and IIO (nonlinear functional response); null models are Null4 and Null5. Dash line on top of the model bar is the level of the explained variance from the model without the environmental variables. 50
4.2 Modularity variance partitioning showing the independent contribution of each of the explanatory variables to the total variance explained for different model combinations. The environmental variables are SE - sampling effort, ND - nearest distance (isolation), AR - area and HPS - human population size (log transformed [log(x+1)]). The band strips shows the 95% confidence interval of the randomly divided variance explained. Bars above this band are important and asterisked bars are significant ($p < 0.1$). Elimination models are IE (linear functional response) and IIE (nonlinear functional response); optimization models are IO (linear functional response) and IIO (nonlinear functional response); null models are Null4 and Null5. Dash line on top of the model bar is the level of the explained variance from the model without the environmental variables. 54

List of Tables

2.1 Types of ecological interactions and the effect each type has on the species involved in the interaction. Effects ranges from positive, negative to neutral outcomes. For example, mutualism has a positve effect on both specie A and species B. 6
2.2 Examples of complex networks, their nodes and links. 7

3.1 Environmental variables: P, A and I are the number of plants, animals, and interactions respectively. C is the connectance (I/PA) ; SE is the sampling effort measured in hours; ND is the nearest distance measured in km; AR is the island area measured in km^2; ME is the maximum elevation of the island measured in km; AG is the age of the island measured in million years (Ma); HPS is human population size. 27
3.2 Nestedness prediction values for our four Adaptive Interaction Switch (AIS) models and two null models, showing the 95% confidence interval $\left(\bar{X} \pm Z_{\alpha/2}\frac{\sigma_{\bar{X}}}{\sqrt{n}}\right)$, where $\bar{X}$ and $\sigma_{\bar{X}}$ are the mean and standard deviation of the model predictions respectively; $Z_{\alpha/2} = 1.96$ is the 95% confidence value; number of samples n is 100 for our AIS models and 1000 for the null models. Elimination models are IE (type I functional response) and IIE (type II functional response); optimization models are IO (type I functional response) and IIO (type II functional response), null models are NULL4 and NULL5. All predictions are significantly different from the observed values ($p < 0.0001$). 36
3.3 Nestedness linear regression results showing the comparison between the observed values and the model predictions, their R squared values, slope and p-values, using Reduced Major Axis (RMA) linear regression. R squared is the coefficient of determination. Elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5. 37

3.4 Modularity prediction values for our four Adaptive Interaction Switch (AIS) models and two null models, showing the 95% confidence interval $\left(\bar{X} \pm Z_{\alpha/2}\frac{\sigma_{\bar{X}}}{\sqrt{n}}\right)$, where $\bar{X}$ and $\sigma_{\bar{X}}$ are the mean and standard deviation of the model predictions respectively; $Z_{\alpha/2} = 1.96$ is the 95% confidence value; number of samples n is 100 for our AIS models and 1000 for the null models. Elimination models are IE (type I functional response) and IIE (type II functional response); optimization models are IO (type I functional response) and IIO (type II functional response), null models are NULL4 and NULL5. All predictions are significantly different from the observed values ($p < 0.0001$), except for the underlined prediction ($p = 0.845$). . . . 41

3.5 Modularity linear regression results showing the comparison between the observed values and the model predictions, their R squared values, slope and p-values, using Reduced Major Axis (RMA) linear regression. R squared is the coefficient of determination. Elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5. 41

4.1 Nestedness Variance Inflation Factor (VIF) showing the VIF values for each combination of the explanatory variables. ND is the nearest distance, SE is the sampling effort, AR is the area, HPS is the human population size (log transformed $[log(x + 1)]$), ME is the maximum elevation, AG is the age. To obtain non collinearity among the variables, we make sure to achieve VIF values less than 5, dropping variables with high VIF values. (a) is the combination of all the environmental variables and the model predictions. (b) is the combination of the covariants without age. (c) is the combination of the covariants without age and maximum elevation. Elimination models are IE (type I, linear functional response) and IIE (type II, non-linear functional response); optimization models are IO (type I, linear functional response) and IIO (type II, non-linear functional response); null models are Null4 and Null5. R squared is the coefficient of determination of the Generalized linear model fittings. 48

4.2 Nestedness multiple regression showing the Generalised linear model (Glm) fitting of the observed values with the model predictions, ND is the nearest distance, SE is the sampling effort, AR is the area, HPS is the human population size (log transformed $[log(x+1)]$). R squared is the coefficient of determination, AIC is the Akaike Information Criterion. Elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5. 49

4.3 Nestedness variance partitioning results showing the percentage each explanatory variable is contributing to the explained variance (R squared) in the observed values. SE is the sampling effort, ND is the nearest distance, AR is the area, HPS is the human population size (log transformed $[log(x+1)]$). Glm is the generalised linear model fitting; elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5. 51

4.4 Modularity Variance Inflation Factor (VIF) showing the VIF values for each combination of the explanatory variables. ND is the nearest distance, SE is the sampling effort, AR is the area, HPS is the human population size (log transformed $[log(x+1)]$), ME is the maximum elevation, AG is the age. To obtain non collinearity among the variables, we make sure to achieve VIF values less than 5, dropping variables with high VIF values. (a) is the combination of all the environmental variables and the model predictions. (b) is the combination of the covariants without age. (c) is the combination of the covariants without age and maximum elevation. Elimination models are IE (type I, linear functional response) and IIE (type II, non-linear functional response); optimization models are IO (type I, linear functional response) and IIO (type II, non-linear functional response); null models are Null4 and Null5. R squared is the coefficient of determination of the Generalized Linear Model fittings. 52

4.5 Modularity multiple regression showing the Generalised linear model (Glm) fitting of the observed values with the model predictions, ND is the nearest distance, SE is the sampling effort, AR is the area, HPS is the human population size (log transformed $[log(x+1)]$). R squared is the coefficient of determination, AIC is the Akaike Information Criterion. Elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5. 53

4.6 Modularity variance partitioning results showing the percentage each explanatory variable is contributing to the explained variance (R squared) in the observed values. SE is the sampling effort, ND is the nearest distance, AR is the area, HPS is the human population size (log transformed $[log(x + 1)]$). Glm is the generalised linear model fitting; elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5. 55

Chapter 1

Introduction

1.1 Background to the research problem

Ecological interactions in the ecosystem are essential, because survival of species in the ecosystem depends on the interactions the species have with other components of the ecosystem. Abiotic components like light, water and nutrients and biotic components like plants and animals, all interact in a unique way, maintaining the functions and services of the ecosystem. For example, pollination and seed dispersal (mutualistic), predation and parasitism (antagonistic), are biotic interactions having strong effect in sustaining the ecosystem. Considering the case of pollination networks, plants and pollinators reciprocally exploit the benefits (e.g. cross-fertilisation and nectar harvest) at certain levels of energetic or nutrient costs (Memmott, 1999), benefiting to the fitness of involved species (Bascompte *et al.* (2003); Bascompte *et al.* (2006)). Plants gain fertilization in the transfer of their pollen grains from their anther to their stigma by pollinators, in return the pollinators benefit from the plants' nectar, fleshy pulp, fragrance etcetera (Memmott, 1999). As such, mutualistic interactions can be important for enhancing and maintaining ecosystem biodiversity (Bascompte and Jordano, 2006).

The representation of ecological interactions in form of networks where species represents the nodes and links the interactions, have been a substantial breakthrough in the understanding of the ecological systems. These networks of interactions are complex, hence mathematical theories and analytical power have enabled useful insight into understanding the ecological and evolutionary processes of species interactions, shaping the ecosystem. With the advent of computers and increased computing power, more data on ecological interactions have being collected, enabling simulations of real communities to be carried out. These advantages have lead to different kinds of model and analysis to be presented by many researchers in explaining the observed network structures.

As a result of species interactions, different network structures emerge in mutualistic networks. Commonly studied network structures that go beyond simple topologies of a network are nestedness and modularity (Fortuna *et al.* (2010); Sebastián-González *et al.* (2014)). Nestedness accounts for how species with few interaction partners (specialists) interact with a subset of species with more partners (generalists) (Bascompte *et al.*, 2003), while modularity explains how species cluster into modules or compartments with species interacting more frequently in modules than expected by chance (Dicks *et al.* (2002); Olesen *et al.* (2007)). These network structures play crucial role in the stability and persistence of mutualistic networks (Okuyama and Holland (2008); Thébault and Fontaine (2010); Staniczenko *et al.* (2010); Allesina and Tang (2012); James *et al.* (2012)). For instance, highly nested structures in mutualistic communities not only can maintain, or rather are associated with, enhanced species richness (Bastolla *et al.* (2009); but see James *et al.* (2012)), but also are robust to habitat loss (Fortuna and Bascompte, 2006) and random extinction of the species (Burgos *et al.*, 2007). Compartmentalised structures minimizes the destabilizing effect in mutualistic communities (Allesina and Tang, 2012). Although complementary relationship is evident between nestedness and modularity (Fortuna *et al.*, 2010), a full understanding of the mechanisms that underlie their emergence is not yet attained.

In explaining the structural emergence of ecological networks, there have being studies based on evolutionary history (Rezende *et al.* (2007); Nuismer *et al.* (2013); Minoarivelo *et al.* (2014)) and ecological fittings of species rewiring (Zhang *et al.* (2011); Suweis *et al.* (2013)). Adaptive rewiring also known as adaptive interaction switch has being reported to occur among species in response to resource availability (Valdovinos *et al.*, 2010). Some evidences have shown that species switch interaction partners in search for resources (Petanidou *et al.*, 2008). Since mutualism is a reciprocal exploitation, rewiring of feeding partners fit the concept of optimal and adaptive foraging theory where species are adaptively changing their diet according to profitability, encounter rate and past experience (Stephens and Krebs (1986); Fossette *et al.* (2011); Zhang and Hui (2014)). Therefore, behavioural adaptation of species in switching interaction partners could potentially be a key process by which nestedness and modularity emerge in a mutualistic community, affecting the community's persistence and stability (Kondoh (2003); Staniczenko *et al.* (2010); Valdovinos *et al.* (2010)).

Ecological communities including Islands harbour a variety of endemic and alien species, with the species richness varying from one island to the other largely due to changes in island size and isolation (MacArthur and Wilson, 1967). Different floras and faunas, found on islands interact among themselves to gain food, support and reproduction, thereby diversifying in abundance. Even as alien species are being introduced into the islands, the aliens involve in

the interactions with native species and endemic species (Traveset *et al.*, 2013), leading to a complex assemblage of the interactions. Island characteristic such as size, isolation, age, elevation can have effect on the network structures (Roura-Pascual *et al.*, 2016). These geographical factors can potentially shape the structures observed on Islands.

Anthropogenic changes consisting of human induced processes affect the composition of species in the community (Sole and Montoya, 2001). Processes like habitat restoration and introduction of alien species can have influence in the functions and services of the ecosystem. These processes can lead to extinction of species, imposing a great threat to ecosystem biodiversity (Sala *et al.*, 2000). Human disturbances/impacts and sampling effort greatly affect community networks, as such they can be important factors determining network structures (Tylianakis *et al.* (2008); Rivera-Hutinel *et al.* (2012); Sebastián-González *et al.* (2014)).

1.2 Problem statement

The maintenance and sustenance of the ecosystem functions and services is crucial as many species are threatened. Loss of interactions of one species can lead to the loss of the other species depending on it for food. Even habitat degeneration can lead to cascading effect on the species. This makes species interaction an important process in the ecosystem that needs to be conserved to ensure the diversity of species. Therefore species interactions are crucial, leading to the emergence of structures that stabilize the community. Our potential hypothesis is in the light of previous network models of species interactions, where adaptive rewiring has predicted the emergence of network structures such as nestedness and modularity in mutualistic and antagonistic networks, leading to stability of the networks (Zhang *et al.* (2011); Suweis *et al.* (2013); Nuwagaba *et al.* (2015)). Concurrently, anthropogenic and geographic factors can have effect in explaining the observed network structures (Rivera-Hutinel *et al.* (2012); Sebastián-González *et al.* (2014); Roura-Pascual *et al.* (2016)).

1.3 Research question

Despite the efforts to comprehend the emergence of network structures, a general concensus on the underlying mechanism is yet unattained. With this problem in mind, we seek to address three research questions:

1. how do models that incorporates adaptive interaction switches explain the emergence of network structures like nestedness and modularity?

2. can geographic factors have effect in accounting for the emergence of network structures?

3. to what extent do anthropogenic disturbance/changes affect or lead to the emergence of network structures?

1.4 Aims and objectives

The structure of a community can alter and foster its stability to perturbations (Okuyama and Holland, 2008). Using the pollination data collected from 10 Galápagos islands, our goal is to understand the emergence of network structures from these ecological communities. To achieve this goal, we want to

1. determine the predictive power of Lotka-Volterra mutualism models that incorporates intra and inter specific competition of species, behaviour of species (adaptive rewiring), linear and non-linear feeding relationships (Hollings type I and type II functional responses), in explaining the observed network structures.

2. determine the effect of geographic factors like island size (area), isolation (nearest distance), age and elevation combined with the model predictions in explaining the observed network structures.

3. determine the effect of human population size and their sampling effort in the emergence of structures.

1.5 Thesis overview

Having introduced the background and stated our problem, highlighted the uprising research questions and explained the main aim of our studies in this Chapter, we now move on to understanding the literature of what has being done and studied in this field by other authors in Chapter 2. Thereafter, in Chapter 3, we give a description of our study system and the development of the Adaptive Interaction Switch (AIS) models. We will then use these AIS models to predict the observed structures by simulations and then we state and discuss the results. In chapter 4, we will investigate the effect of environmental variables (geographic and anthropogenic factors) in explaining the observed structures. Statistically, we will analyse the resulting data and discuss the results. Lastly, we conclude in Chapter 5, giving a summary of the work done, limitations in this study and further extensions of this study.

Chapter 2

Literature Review

In the 19th century, appreciable works have been done by many authors starting from the earliest works of Lotka-Voltera, May, MacArthur, Erdos and Renyi, Hollings, to the latest ones. These works serve as the foundation upon which we build our theoretical studies in this thesis. In this Chapter, we will give a review of what is know and has been done by other researchers as well as defining some key terms of our focus. We start with a section on the description of ecological interactions that yields complex networks. Thereafter we give details and properties of complex networks and briefly describe some network models. From there, we look into the structures of complex networks, focusing more on the structures of mutualistic networks. Dynamical models incorporating functional responses that have been used before will be highlighted more importantly on adaptive rewiring of species interactions. Since we are looking at the network structures of Galápagos islands, we will briefly explain island biogeography and relate it with other environmental variables we are considering. Last but not the least, we will explore the measures of stability of network structures.

2.1 Ecological Interactions

Ecological interactions are prevalent in the ecosystem. They occur between and within the biotic and abiotic components of the ecosystem, persisting for a long time. Species in the ecosystem engage in interactions directly or indirectly to sustain their biodiversity and maintain their functions in the ecosystem (Newman *et al.*, 2011). Food, shelter, reproduction among other functions of ecosystem are obtained through the interactions of species. Most interactions can be on a long term basis, sustaining species richness and stabilizing the community. Unlike non-interacting studies to ecology (Alonso *et al.*, 2006), species in the ecosystem interact in maintaining the ecosystem biodversity.

Species interactions can also be seen as energy flow from one species to the

other, like in food webs. There are two kinds of interactions namely: intraspecific and interspecific interactions in the ecosystem, the former is the interaction between individuals of the same species while the latter is the interaction between individuals of different species. A species interaction could have a positive, negative or neutral effect on the other partner species (Table 2.1). Based on this kind of effect, we can divide ecological interactions into five types: Mutualistic, Antagonistic, Competition, Commensalism and Ammensalism interactions (Figs. 2.1 and 2.2).

Table 2.1: Types of ecological interactions and the effect each type has on the species involved in the interaction. Effects ranges from positive, negative to neutral outcomes. For example, mutualism has a positve effect on both specie A and species B.

Ecological Interaction type	Species A	Species B
Mutualism	Positive	Positive
Antagonistic	Negative	Positive
Competition	Negative	Negative
Commensalism	Neutral	Positive
Ammensalism	Neutral	Negative

Species interactions can be one way (unipartite) or two ways (bipartite) in nature. More detail will be discussed later.

Mutualistic interactions is a two way interaction in the ecosystem, having beneficial effects on the species involved, for example pollination and seed dispersal (Fig. 2.1). Studies have shown that mutualistic interactions foster biodiversity (Bastolla *et al.*, 2009), because most of tropical plants (almost 90% of plants) depend on animals for their seed dispersal (Jordano *et al.*, 2000). Examples of pollinators are hoverflies, solitary bees, bumble bees, bats or birds (Memmott, 1999) (Fig. 2.1).

Ecological interactions have given rise to complex networks, so in the next section, we will describe the emerging complex networks.

2.2 Complex Networks

The way in which species in a community interact can be likened to a network where the nodes of the network are the different species and the links of the network are the species interactions. Many species and their interactions have given rise to hundreds and thousands of nodes and links making the network complex. Complex networks not only occur in ecological interactions but are also evident in social interactions, internet, World Wide Web, physical and many other biological interactions (Table 2.2). Therefore a network approach

(a) Honey bee *(b) Hoverfly* *(c) Bumblebee*

(d) Solitory bee *(e) Pollinator bird* *(f) Black bird*

Figure 2.1: Mutualistic interactions between plants and animals. (a)-(e) are pollination and (f) is seed dispersal (Severns (2016); Cole (2016); Tribune (2016); Batra (2016); Tutorvista (2016); Moore (2001)).

to studying complex system is fundamental in uncovering the uprising features of the system and their properties. Different studies of complex networks have showed common features attributed to these complex networks (Dunne *et al.* (2002); Newman *et al.* (2011)).

Table 2.2: Examples of complex networks, their nodes and links.

Complex networks	Nodes	Links
Food webs	Species	Ecological interactions
World Wide Web	Web pages/HTML document	Hyper-links
Internet	Routers/Computers	Physical or Wireless links
Cell	Chemicals	Chemical interactions
Social network	Human beings	Social relationships
Nervous system	Nerve cells	Axons
Social science	Individual	Social interactions
Scientific publication	Papers	Citations to the paper

The easiest form of complex network are random graphs introduced by two mathematicians Paul Erdos and Alfred Renyi in 1959, nevertheless, the theory

(a) Intraspecies competition

(b) Interspecies competition

(c) Plant herbivory

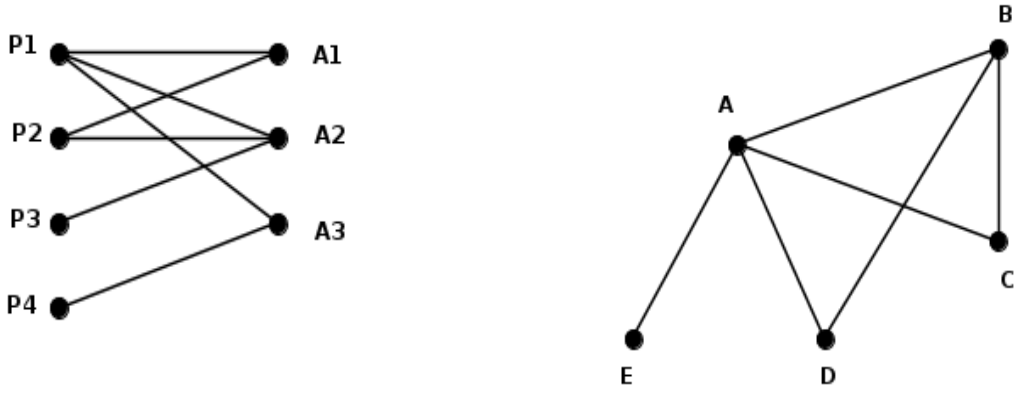
(d) Predator Prey

Figure 2.2: Ecological interactions: (a) and (b) are competition; (c) and (d) are antagonistic interactions (Mohammed (2016); Manaster (2016); Bar-Yam (2016a); Bar-Yam (2016b)).

of graphs was introduced in the 18th century by Leonhard Euler. We will now define graphs and its elements.

2.2.1 Graph

A graph is an ordered pair of set $G = \{N, E\}$, where N is the set of nodes or vertices of the graph and E is the set of edges or links that connects two nodes in the graph. For example, Fig. 2.3b has 5 nodes and 6 edges. Here, we also refer to graphs as networks. For instance, in ecological network, the nodes are the species and the links are the interactions.



(a) Bipartite graph *(b) Unipartite graph*

Figure 2.3: Graph types

2.2.1.1 Types of graph

Based on the mode of interaction of species in a network, a graph can be unipartite or bipartite (Fig. 2.3). In a unipartite graph, interactions are among the species (Fig 2.3b) just like in food webs but in a bipartite graph, species interact between not within themselves (Fig 2.3a). This makes plant-animal mutualistic network to be described as a bipartite graph. For example plant-pollinator, plant seed-dispersal, plant-ant (Fig. 2.1) are mutualistic communities that have being represented using bipartite graphs (Jordano *et al.* (2003); Bascompte *et al.* (2003); Bascompte *et al.* (2006)). Irrespective of the type of interaction, common properties are attributed to complex networks.

2.2.2 Properties of complex networks

Complex networks exhibit certain properties with respect to their form and shape. We will briefly explain four of the properties namely: small worlds, clustering, degree distribution and resilience.

2.2.2.1 Small worlds

This property describes the distance between two nodes in a complex network. This distance is a short path length, equivalent to the number of links between the two nodes. Although the size of complex networks can be large, they still have a small world property (Milgram, 1967). In a network with no direction (e.g Fig. 2.3b), the mean shortest distance between pairs of nodes in the network is given by

$$l = \frac{2}{N(N+1)} \sum_{i \geq j} d_{ij} \,, \tag{2.2.1}$$

where d_{ij} is the shortest distance from node i to j and N is the number of vertices (Latora and Marchiori, 2001).

2.2.2.2 Clustering

Watts and Strogatz (1998) came up with the concept of clustering coefficient which is the tendency of nodes to cluster in a complex network. This is given by

$$C_i = \frac{2E_i}{k_i(k_i - 1)}\,, \tag{2.2.2}$$

where k_i represents the number of links to node i, and E_i is the links that exists between k_i nodes. For the whole network, the clustering coefficient C is given by the average of all each C_i. In random graph, the clustering coefficient is the same as the probability of generating the graph.

2.2.2.3 Degree distribution

The nodes in a complex network do not have the same degree (that is number of links to a node). These nodes can be characterised by a probability distribution function $P(k)$ that a node selected at random has k links. Random graphs have poisson distribution for its degree, but other complex networks have power-law degree distributions, that is

$$P(k) \sim k^{-\gamma}\,, \tag{2.2.3}$$

where k is the degree and γ is the degree exponent, usually in the range $2 < \gamma < 3$. Networks having power-law degree distributions are scale-free (Barabási and Albert, 1999).

2.2.2.4 Network resilience

This property is a network's resilience to the removal of a node. Albert *et al.* (2000) examined the effect of the removal of nodes from internet and World Wide Web networks. These networks were highly resistant to the removal of nodes with few links but became vulnerable to loss of other nodes when nodes with many links were removed.

With these network properties in mind, certain network models have being put forward in understanding complex networks. We now consider some network models that was studied before.

2.2.3 Network models

Over the years, network modelling have progressed from the simplest form of complex networks - random graphs to models like small world models and scale free models of complex networks.

2.2.3.1 Random network models

The first model attempt to represent random network was by Solomonoff and Rapoport (1951). They developed a *random net* model (Rapoport, 1957). Years later, Erdös and Rényi (1959) came up with a model called *random graph* (Erdös and Rényi (1959); Erdös and Rényi (1960); Erdös and Rényi (1961)). Random graphs are the easiest form of complex networks. This random graph is generated with a probability

$$p = \frac{N(N-1)}{2}, \tag{2.2.4}$$

that given N nodes of a graph, any two nodes are connected by a link or an edge. This simple construction of a random graph has posed questions, to determine if complex networks like food webs, follow the same random pattern of network generation by Erdös and Rényi. Some other mathematicians also studied random graphs (Karoński (1982); Bollobás (1998)) and from their studies, many exact and approximate results have being obtained. Complex networks have their unique patterns but if they deviate from random graphs, then, there is a need to quantify the observed pattern by developing other tools. In recent years, networks from various systems are observed to be non-random in nature, therefore mechanisms that account for their pattern are being investigated.

2.2.3.2 Small world models

Motivated by the high clustering of nodes and small average shortest path length in a complex network, Watts and Strogatz (1998) proposed a small world model. They achieved this by starting with a network of N nodes (for example, a ring lattice) where all the nodes are connected to their nearest k neigbours, having $\frac{k}{2}$ links on either side of the node, given that $N \gg k \gg ln(N) \gg 1$. Then with a certain probability p, they randomly rewire each link in the network in such a way that a node connected to itself and repetition of links are avoided.

Erdös and Rényi random graphs are limited in the sense that they give low clustering coefficient, in contrast, Watts and Strogatz model show high clustering coefficient as seen in real networks.

2.2.3.3 Scale free models

Scale free models are due to the power law degree distribution of complex networks. In an attempt to capture the network dynamics rather than focusing on network topologies only as in the case of Erdös and Rényi (ER) and Watts and Strogatz (WS) models, Barabási and Albert (1999) developed a scale free network model. Their success was based on two limitations in ER and WS

models. First is that ER and WS models has a fixed number of nodes N, which is not the case of real networks that grows through the addition of a new node. Secondly, ER and WS models assumes that the probability that a node is connected does not depend on the degree of the node, which does not depict real network, because most real networks show a preferential attachment in the way a node is connected. Motivated by these two mechanisms, growth and preferential attachment, Barabasi and Albert (BA) developed scale free model with power-law degree distribution.

The BA model begins with a small number of nodes n_0, then at every time step, a new node is added to the system. This new node connects to other nodes through n links where $n \leq n_0$. The probability that the new node will be connected to a certain node n_i depends on the degree k_i of the node. This probability is given by

$$\prod(k_i) = \frac{k_i}{\sum_j k_j}\,. \tag{2.2.5}$$

Therefore, this will result to a network with $n_0 + t$ nodes and nt links after t time steps.

The way nodes and links are organized in a complex network give rise to different structures which we will go through in the next section.

2.3 Structures of Complex Networks

Species interact in a structured way (Bascompte *et al.* (2003); Montoya *et al.* (2006); Pascual and Dunne (2006); Okuyama and Holland (2008); Bastolla *et al.* (2009)) unlike the random pattern reported by May (1973) and Chesson (2000). For example, the structure of social network can affect the spread of informations on it. Structures like degree distribution, network size, nestedness and modularity are evident in most complex networks. Now we choose to narrow down to the structures pertaining to ecological networks, most especially mutualistic networks.

2.3.1 Structures of Mutualistic Networks

Mutualistic interactions between plants and animals have generally evolved different structures in the community. As plants provide resources like pollens, floral parts, nectar, fragrances and fleshy pulp to animals when they interact, so also animals like hoverflies, solitary bees, bumble bees, bats or birds (Memmott, 1999); Fig. 2.1) assist in moving reproduction gametes from the male to the female part of the plant (Proctor *et al.* (1996); Jordano (1987)). Recent work has shown how network of mutualistic interactions coevolve (Memmott (1999); Jordano *et al.* (2003); Bascompte *et al.* (2003); Vázquez and Aizen

(2004)). These studies are important in comprehending the coevolution and persistence of species assemblages. Mutualistic networks have shaped biodiversity (Ehrlich and Raven (1964); Thompson (1994)) and so, we now dive into how structural properties influences mutualistic communitites.

2.3.1.1 Community size

Community size is the total number of species in a community. It is also know as species diversity. Jordano (1987) conducted a quantitative study by assembling 36 pollination and 19 seed dispersal networks from published sources and described the distribution of interaction between the mutualistic species. From his study, he found that number of interactions increased faster in seed dispersal than in pollination networks with increase in size of the networks. Also, as the number of mutualistic species increases, the number of interaction increases but connectance decreases exponentially. Bastolla *et al.* (2009) quantified the effect of network structure on the number of species that stably coexist in the community. Although May (1973) proposed that community size has a negative relationship on stability of foodwebs, Okuyama and Holland (2008) showed that increases in community size increased the stability of the community.

2.3.1.2 Connectance

Connectance is the proportion of realised links or interactions in a network. This is given by

$$C = \frac{I}{PA}, \tag{2.3.1}$$

where I, P and A are the numbers of interactions, plants and animals respectively. For example the connectance of Genovesa network (Fig. 2.4) is 0.11, having 16 plants, 37 animals and 64 interactions. This structure was studied by Jordano (1987). In his work, connectance decreased as species were added to a network. In other words, connectance decreased with size although the number of interactions increased with size. Olesen *et al.* (2007) found that connectance increased with nestedness and decreased with modularity.

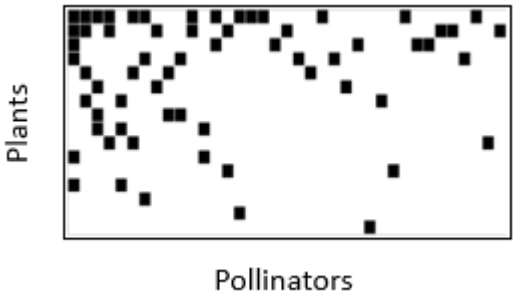


Figure 2.4: Genovesa island network

2.3.1.3 Strength and asymmetry

Strength or dependence of a network is the value of the extent to which plants and animals depend on each other for food and nourishment, that is the relative interaction strength between the plants and the animals (Jordano, 1987). For example, percentage of visits recorded or percentage of seed ingested by the plants and animals, including the reproductive effect of the species (Jordano (1987); Vázquez *et al.* (2005*a*)). Jordano (1987) found that dependence decreased with increase in size of mutualist partners.

Studies have shown that mutualistic communities have many weak intearactions and few strong interactions (Memmott (1999); Bascompte *et al.* (2006); Vázquez *et al.* (2007)). Therefore, mutualistic networks are asymmetric, facilitating biodversity and can improve long term coexistence of the species (Bascompte *et al.*, 2006). By investigating the effect of the weak dependencies, Bascompte *et al.* (2006) compiled 19 plant-pollinator and 7 plant-frugivores networks from artic and tropical communities, to understand the organisation of these network and how these networks coexist. They calculated an asymmetry index, which represents the relative variants observed between two mutual dependencies using

$$AS(i,j) = \frac{d^P_{i,j} - d^A_{j,i}}{\max(d^P_{i,j}, d^A_{j,i})}, \tag{2.3.2}$$

where $d^P_{i,j}$ is the dependence of plant i on animal j, $d^A_{j,i}$ is the dependence of animal j on plant i and $\max(d^P_{i,j}, d^A_{j,i})$ is the maximum between the two dependence values. From their results, the frequency distribution of dependences were skewed to the right for both pollination and seed dispersals networks, having a few strong dependencies and so many weak ones. This is in accordance with the work of Jordano (1987). Generally, Bascompte *et al.* (2006) advocated for a weak dependence in species interaction, to achieve stable coexistence between the mutualistic species.

2.3.1.4 Degree distribution

The degree of a species in a network is the number of links or interactions it has with other species in the network. Therefore, the frequency distribution of the degrees of all species in the network is the degree distribution of the network. Mutualistic networks have being shown to have a power law degree distribution given by

$$P(k) \sim k^{-\gamma}, \quad 2 < \gamma < 3 \tag{2.3.3}$$

where k is the degree and γ is the degree exponent which increases with connectivity of a network (Okuyama and Holland, 2008). Species strength is the quantitative extension of species degree. Species can have equal degree but various strengths of interactions (Bascompte *et al.*, 2006).

2.3.1.5 Nestedness

Mutualistic networks have being found to be nested (Bascompte *et al.* (2003); Guimarães *et al.* (2006); Guimarães *et al.* (2007); Lewinsohn *et al.* (2006); Ollerton *et al.* (2007)). The word nestedness was adopted from island biogeography to shed more light on how species are redistributed among a set of islands (Atmar and Patterson, 1993), by thinking of a plant as an island that contains several animals on it. Nestedness in a mutualistic network is when specialists species interact with a proper subset of species that interacts with generalists (e.g Fig 2.5a). In order words, nestedness moves from species that interacts with one species, and gradually attain species that interacts with all species (Bascompte and Jordano, 2006).

By analysing 27 plant-frugivores and 25 plant-pollinators network, Bascompte *et al.* (2003) showed that mutualistic networks are highly nested. To explain nested structures, studies have shown that nestedness emerge as a result of species abundance (Vázquez *et al.*, 2005*b*), geographic coevolution (Thompson, 2005), past evolutionary history (Minoarivelo *et al.*, 2014). Bastolla *et al.* (2009) has shown that nestedness reduces competition between the interacting species and that nestedness will emerge if a new species entering a community where competition is small, interacts with generalist species. Although some studies have shown the destabilizing effect of nestedness (Allesina and Tang (2012); James *et al.* (2012)), nevertheless, mutualistic communities with high nested structures gradually evolve, not only increasing the species biodiversity (Bastolla *et al.*, 2009) but also stabilizing the community (Okuyama and Holland (2008); Thébault and Fontaine (2010); Staniczenko *et al.* (2010)). Nogales *et al.* (2015) found high nestedness in the seed dispersal networks of Canary islands, which prevented community disassembly, more importantly, extinction of rare species.

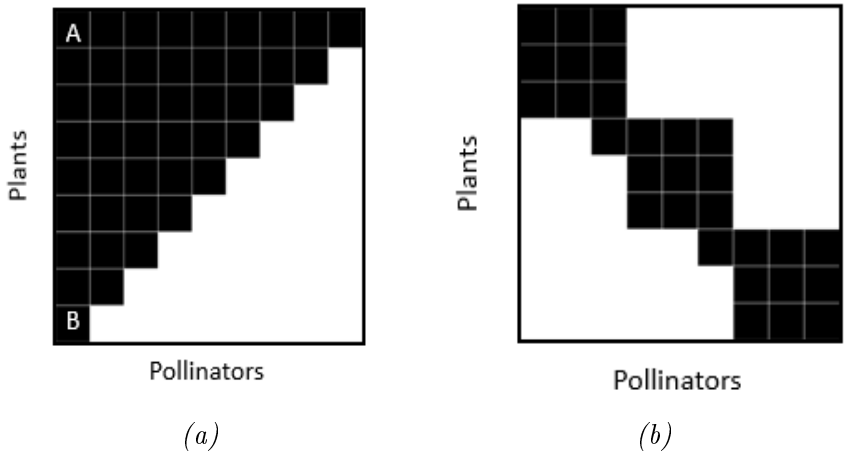
Almeida-Neto *et al.* (2008) developed a metric for computing nestedness of a network, based on decreasing fill (DF) and paired overlap (PO) of the entries of the network. Representing the network in form of a binary matrix, DF can either be 0 or 100 depending on the marginal total (MT) of the rows or columns of the matrix. That is for rows i, j in the matrix,

$$DF_{ij} = \begin{cases} 1 & \text{if} \quad MT_j < MT_i \\ 0 & \text{if} \quad MT_j \geq MT_i \end{cases}, \tag{2.3.4}$$

and PO_{ij} in the rows i, j is the percentage of 1's in row j that overlaps with 1's in row i. Same applies for the columns of the matrix. Therefore the metric is given by

$$NODF = \frac{\sum_{i<j} N_{paired}}{\left[\frac{P(P-1)}{2}\right] + \left[\frac{A(A-1)}{2}\right]}, \tag{2.3.5}$$

where NODF is Nestedness metric based on overlap and decreasing fill; $N_{paired} = DF_{ij} \times PO_{ij}$ for all the combinations of rows P and columns A. NODF value can either be 0 (no nestedness) or 1 (perfect nestedness). For example, Figs. 2.5a and 2.4 have nestedness values of 1 and 0.17 respectively.



(a) *(b)*

Figure 2.5: Network structures representing nestedness and modularity, where each box represents an interaction between plants and pollinators. (a) is a perfectly nested network, having interaction A which is between a generalist plant and a generalist pollinator, and interaction B, between a specialist plant and a generalist pollinator. (b) is a compartmentalised network.

2.3.1.6 Modularity

Modularity explains how species interact more frequently in modules or compartment than expected by chance. These modules are link dense regions in a network (e.g. Fig. 2.5b). This means that modularity is the extent to which

species interaction are organised into modules, where a species is more linked closely to the species in the same module with it, than with species in other modules of a network.

Paine (1963) and May (1972) searched for compartments in food webs and they argued on the effect of compartmentalisation in the stability of food webs. Paine (1963) noted compartments stabilizes a food web community but this was opposed by May (1972), that compartmentalized structures destabilize the community. Jordano (1987) observed 7 assemblages in his study of 55 networks, noting that compartments increased with size. In the analysis of 51 networks by Olesen *et al.* (2007), networks more than 150 species were modular, showing that modularity increases with the size of a network. Recently, Nogales *et al.* (2015) explored the network structure of modularity on seed dispersal networks of Galápagos islands. In their study, high modularity was reported for Galápagos networks, minimizing the risk of extinction cascades.

Newman and Girvan (2004) provided a measure for the modularity of a network, given by

$$M = \sum_{s=1}^{N_M} \left[\frac{l_s}{L} - \left(\frac{d_s}{2L} \right)^2 \right], \tag{2.3.6}$$

where N_M is the number of modules in the network; L is the number of links in the network; l_s is the number of links in a module s; d_s is the sum of the degree of the nodes in module s. To look for partitions with high modularity, Guimera and Amaral (2005) used simulated annealing, a stochatic optimization algorithm in getting the best determination of modules. This they did by directly maximizing equation (2.3.6). Simulated annealing allowed for low cost in the computation of modularity.

Having considered the structures of mutualistic networks, we now move on to dynamical models describing species interaction in the next section.

2.4 Dynamic Network Models

Mutualistic communities are dynamic systems where species frequently enter and exit leading to the establishment of interactions, extinctions and varying densities of species. Models describing changes in population densities of the interacting species have being put forward (May, 1981). This moves beyond simple exponential growth model

$$\frac{dN}{dt} = rN \tag{2.4.1}$$

to density dependent logistic models

$$\frac{dN}{dt} = r\left(1 - \frac{N}{K}\right)N\,, \tag{2.4.2}$$

where N is the population density of the species, r is the per capita growth rate, K is the carrying capacity and $\frac{dN}{dt}$ is the rate of change of population densities.

By modifying equation (2.4.2), changes in densities of mutualistic species have being represented through Lotka-Voltera competition model, including the mutual benefits obtained from the interaction (functional response), that is

$$\frac{dN_i}{dt} = r_i\left(1 - \sum_{j=1}^{n}\alpha_{ij}N_j\right)N_i + f(N_j)\,, \tag{2.4.3}$$

where N_i is the density of species i, r_i is the per capita growth rate of species i in the absence of competition and mutualism, N_j is the densities of partner species, α_{ij} is the competition term and $f(N_j)$ is the saturation to the growth rate of i by the densities of j, that is the functional response.

We now describe models with functional response and because species are adaptively changing their diets due to resource availability, models incorporating adaptive rewiring will be discussed.

2.4.1 Functional Responses

Functional response depicts the per-capita growth rate of one population varying as a function of the density of another population. There are three types of functional responses namely: Type I, II and III (Holling, 1959). Type I is a linear functional response denoted by

$$f(N) = \beta N\,, \tag{2.4.4}$$

while type II and III are nonlinear functional responses given by

$$f(N) = \frac{\beta N^k}{1 + h\beta N^k}\,, \tag{2.4.5}$$

where N is the species density, β represents the mutualistic benefits obtained from the interaction, h denotes the handling time, k is an exponent that differentiates type II ($k = 1$) from type III ($k > 1$) (Fig. 2.6).

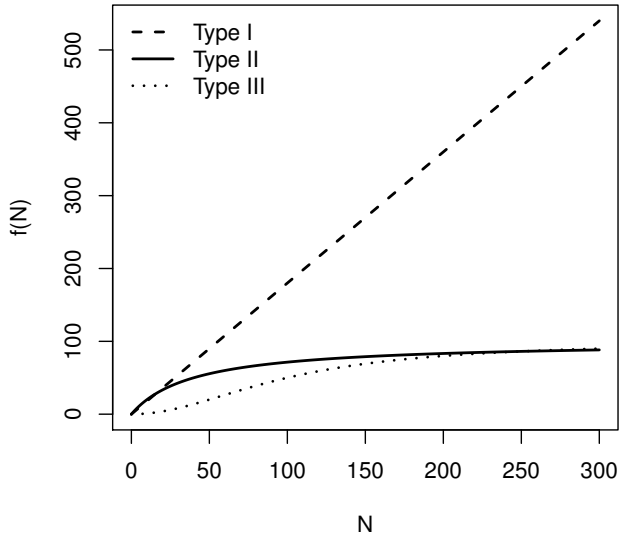


Figure 2.6: Types of functional responses.

Linear functional response in mutualism models is ever increasing mutual benefits with increasing population sizes, leading to unbounded population growth. In large densities of species, linear functional response may not be suitable in representing the per capita growth rates of species. Therefore, nonlinear functional response play a realistic role here, not only limiting the maximum growth rate, but preventing divergence in large densities of species (that is, the denominator of equation (2.4.5)) (Rosenzweig and MacArthur (1963); Tilman (1982); Holland *et al.* (2002)).

Bascompte *et al.* (2006) advocated for weak interaction strength in models with linear functional response, for stable coexistence of species. In contrast, Holland *et al.* (2006) argued that model with non-linear functional response does not need weak interaction strength for coexistence of the species. They noted that, if the handling time and all other parameters are positive, there will be equilibrium in the densities of the species. Models with linear functional response assumes weak interactions for stability (Bascompte *et al.*, 2006), otherwise the continuous positive feedback of mutualism makes the community unstable as in case of May (1973)'s model which categorized mutualistic networks as destabilizing. By adding nonlinear functional response to the model of mutualistic networks, the community becomes stable (Holland *et al.* (2006); Okuyama and Holland (2008); Bastolla *et al.* (2009)). Non-linear functional response incorporates a real feature of a species abundance and how the abundance of other mutualistis species saturates the increase in the species abun-

dance (Holland *et al.*, 2002). Hence, mutualistic communities with non-linear functional response always have stable steady states while those with linear functional response has local stable steady states if the interactions between species are weak or aymmetric.

2.4.2 Adaptive Interaction Switch (AIS) models

Due to resource availability and environmental pertubations, mutualistic species switch their interaction partners. For example, pollinators switch the plants they interact with, as in Whittall and Hodges (2007). The structure of pollination network change in a continuous manner as a result of resource availability to pollinators and plants, giving rise to a behavioural switching of resources and stabilizing communities. For instance switching stabilized food webs (Kondoh, 2003).

Zhang *et al.* (2011) used an adaptive interaction switch model to predict nestedness in 81 mutualistic real networks. This switching algorithm focused on a species fitness by allowing the species to switch to highly beneficial species and discard less beneficial ones. Moreover, it can be likened to Alfred Russel Wallace's natural selection through the elimiantion of unfit species. Concurrently, as less beneficial species are eliminated, new species are being tried out. Nestedness was robustly predicted through this switching algorithm, accounting for 73% variation in the observed structure. Also, in the 61 antagonist real networks studied by Nuwagaba *et al.* (2015), an adaptive rewiring and random drift model (just as in Zhang *et al.* (2011)) was used to predict network structures of nestedness, modularity and node-degree distribution. The hybrid rule of adaptation plus random drift accounted for more than 90% and 50% variations in the observed modularity and nestedness respectively. In fact, more than 98% of the real networks had no significant difference between the observed and predicted node degree distribution.

An optimization based interaction switch was proposed by Suweis *et al.* (2013). In their analytical analysis, nested structures emerged through the maximization of species richness at equilibrium in a mutualistic community. In addition, their analytic and numerical techniques showed that an increase in the abundance of a species leads to increase in the total number of species and hence increase in the nestedness of the network.

In the networks that Jordano (1987) studied, he noted that distribution of interactions might depend on species fitness within the compartments, and the effecting of fitness can determine the intensity of interaction and diversity. Therefore species can switch to maximize its fitness gain. Rewiring is not only shifts in interactions but also in intensity of interactions (Tylianakis *et al.*, 2007). Although Gilljam *et al.* (2015) argued the negative effect of species

rewiring by using 7 natural and theoretical food webs, where they found that rewiring is more likely to aggravate the effect of species loss. Nevertheless, Kaiser-Bunbury *et al.* (2010) noted that rewiring improved the robustness of networks to species extinction.

Having discussed the dynamical models of changes in species density and models incorporating species rewiring, we now move on to environmental variables that can affect the network structures in one way or the other.

2.5 Environmental Variables

The theory of island biogeography has helped in investigating the factors that affect the diversity of species on islands. Island size (area) and isolation (distance from the mainland) as identified by MacArthur and Wilson (1967) have played a prominent role in determining the number of species on islands. Size affects the extinction rate of species, meaning that large islands have low rate of extinction and small islands have high rate of extinction, this can be as a result of competition for resources. Isolation affects immigration rate, that is close islands have high rate of immigration and far islands have low rate of immigration. Therefore, immigration and extinction rates affects the species richness, the former increases the species richness through the introduction of new species on the islands while the latter reduces species richness. Also, high species richness on an island can reduce immigration rate and increase extinction rate. Hence, stability in the species richness is attained when the immigration rate equals extinction rate (Fig. 2.7).

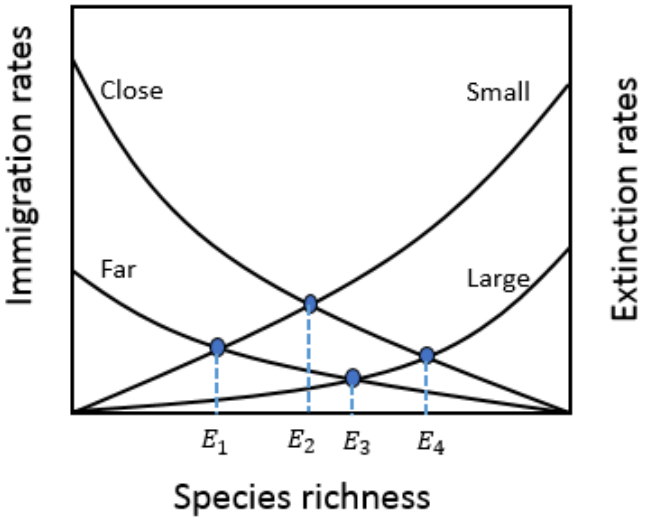


Figure 2.7: Island biogeography theory showing the immigration and extinction rates with species richness. E_1 is the equilibrium point for species richness in far and small islands, E_2 is the equilibrium point for species richness in close and small islands, E_3 is the equilibrium point for species richness in far and large islands and E_4 is the equilibrium point for species richness in close and large islands.

Roura-Pascual *et al.* (2016) studied the distribution and diversity of ants on 102 islands. They found that island area and distance individually accounted for more than 50% variation in native ant richness. Trøjelsgaard *et al.* (2013) found an humped shaped pattern in the relationship between plant species richness and island age from the Canary islands. Studying the community assembly on Hawaiian islands, Rominger *et al.* (2015) found modularity of the networks to increase with age while nestedness decreased with age of the islands.

Rivera-Hutinel *et al.* (2012) investigated the effect of sampling effort on network structures using 10 plant-pollinator networks, comprising of 186 plants and 336 pollinators. From their analysis, network structures were affected by sampling efforts. Sebastián-González *et al.* (2014) tested for the effect of humans in network structure and they found that areas with high human impact were more nestedness and less modular.

In the next section, we now consider how stability of mutualistic networks are measured.

2.6 Measures of Stability

A major focus on ecological networks is their stability. Over several years, empirical and theoretical studies have focused on mechanisms to stabilize a community. Kaiser-Bunbury *et al.* (2010) investigated the stability of two mutualistic networks taking into account their behavioural shifts in the communities. According to Bascompte and Jordano (2006), there are two stability approach in community dynamics:

1. Demographic stability: revolves round a local stability with the steady states. Describing little perturbations around the steady state will increase or decrease in time (May, 1972).
2. Topological stability: This approach checks for stability of a network by the removal of nodes (Pimm (1979); Paine (1980); Amaral *et al.* (2000)).

Comprehending how ecological communities respond to pertubations is crucial, because the perturbations can either amplify or die out. Due to the complex nature of ecological communities, several measures of stability are used including asymptotic stability, resilience, persistence, robustness, resistance to invasions among others. (Pimm (1984); McCann (2000); Ives and Carpenter (2007); Allesina and Tang (2012); Donohue *et al.* (2013)). We now give a brief description of some of these measures of stability.

2.6.1 Asymptotic Stability

This is the earliest measure of ecological stability. It is measured by the maximum real part of the eigenvalue ($max(Re(\lambda))$) of a Jacobian matrix or interaction matrix at the equilibrium densities. If $max(Re(\lambda))$ is negative, then the system is stable, meaning that small perturbations from the steady states will dampen, converging back to the equilibrium density. Otherwise if $max(Re(\lambda))$ is positive, the system is unstable, meaning that there will be amplification in any small perturbation from equilibrium, which can lead to loss of species in the community. (May (1973); Allesina and Tang (2012); Feng and Takemoto (2014)).

2.6.2 Resilience

If the maximum real part of the eigenvalue ($max(Re(\lambda))$) of a Jacobian matrix is negative, then the absolute value, $|max(Re(\lambda))|$ gives the resilience of the matrix or network. The time taken for a network to recover from perturbation is given by the inverse $\frac{1}{|max(Re(\lambda))|}$. Therefore systems that return to equilibrium on time are said to be resilient. (Thébault and Fontaine (2010); Okuyama and Holland (2008)).

2.6.3 Persistence

Persistence is the total number of coexisting species that did not go extinct at the equilibrium densities in the system. (Bascompte *et al.* (2006); Bastolla *et al.* (2009); Thébault and Fontaine (2010)).

2.6.4 Robustness

Measured in a system as its ability to resist extinction cascades when a target species is removed from the system (Kaiser-Bunbury *et al.*, 2010).

Chapter 3

Adaptive Interaction Switch Models and Study system

Adaptive Interaction Switch (AIS) models also known as adaptive rewiring models have being reported to predict the emergence of network structures as we have discussed in Chapter 2. By using the data collected from Galápagos Islands, we will now predict the emergence of network structures from these data. We start off this Chapter by giving a description of our study system and then we will describe the development and outcomes of the models.

3.1 Galápagos Islands/Study data

The Galápagos islands are located on the equator at the eastern pacific ocean (with the coordinate 1°40′N-1°36′S, 89°16′-92°01′W), 960 km to the west of Ecuador in South America (Fig. 3.1). There are 18 main islands and many islets. Española island is the oldest, about 3-3.5 Million years ago (Ma) and the most southern island, having about 163 km average distance from the other main islands. The youngest island is Fernandina, which erupted about 0.035-0.07 Ma (Poulakakis *et al.*, 2012). This group of islands (archipelago) consist of 7880 km^2 of land, distributed across 45 000 km^2 of ocean. The largest Island is Isabela having an area of 4670 km^2 with an elevation of 1.707 km. About 20 000 humans live in some of the islands which includes Floreana, Isabela, San Cristóbal and Santa Cruz islands (Conservancy, 2016).

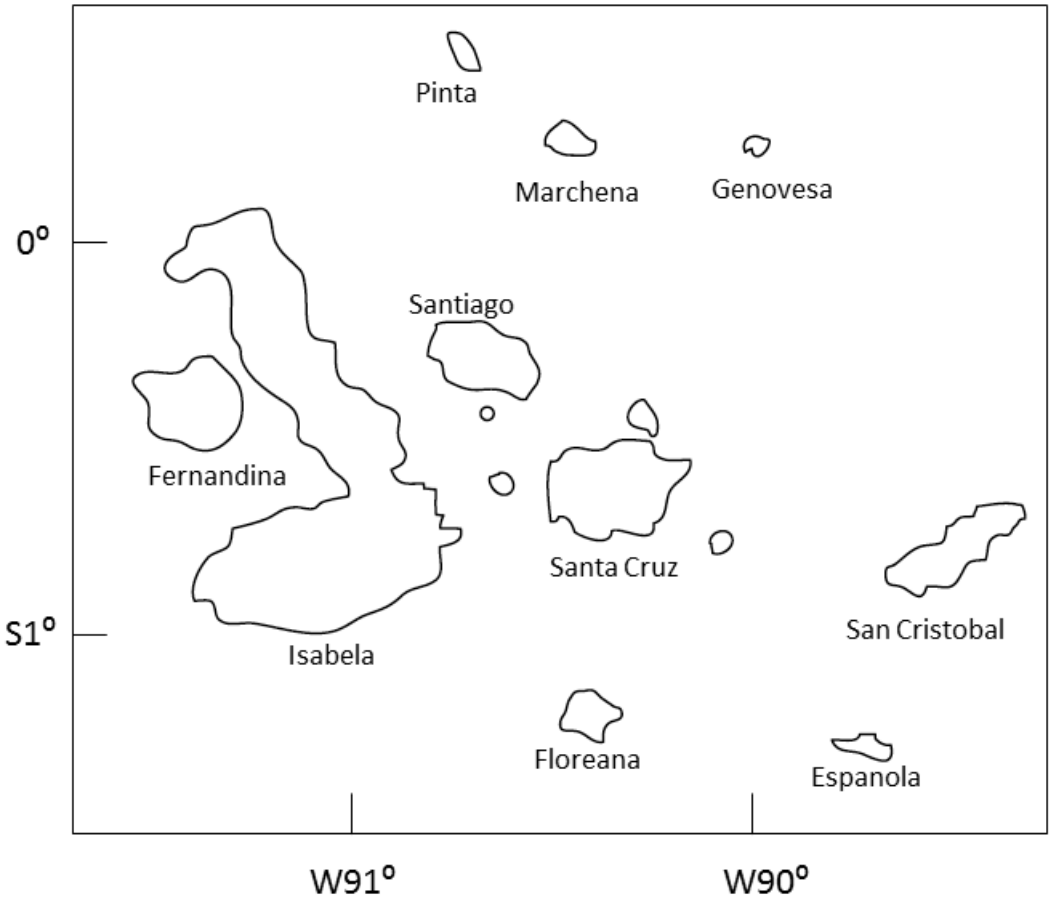


Figure 3.1: Galapagos Islands.

The Galápagos flora consist of more than 1400 species of which 59% are aliens, 14% are endemic and 27% native vascular-plants (Jaramillo *et al.*, 2012). Species richness of the terrestial invertebrates constitutes 51% of the total biodiversity (Herrera and Roque-Álbelo, 2012). Pollination data of visitations was collected from a systematic survey of recording pollination events on 10 islands, including Española, Fernandina, Floreana, Genovesa, Isabela, Marchena, Pinta, San Cristóbal, Santa Cruz and Santiago islands (Traveset *et al.*, 2013). For each island, visitation interactions to flowering plants were sampled in February 2010 and 2011. In these samplings, the flowering plants were observed in a random way, classifying all flower-visiting animals as pollinators (Traveset *et al.*, 2013). For each island, sampling efforts ranged from 6-109 hours (Table 3.1).

Table 3.1: Environmental variables: P, A and I are the number of plants, animals, and interactions respectively. C is the connectance (I/PA) ; SE is the sampling effort measured in hours; ND is the nearest distance measured in km; AR is the island area measured in km^2; ME is the maximum elevation of the island measured in km; AG is the age of the island measured in million years (Ma); HPS is human population size.

Islands	P	A	I	C	SE (hrs)	ND (km)	AR (km^2)	ME (km)	AG (Ma)	HPS
Espanola	11	16	31	0.18	16.75	162.58	60	0.21	3.50	0
Fernandina	18	60	125	0.12	70.25	128.00	642	1.48	0.07	0
Floreana	11	4	13	0.30	6.30	125.85	173	0.64	2.30	100
Genovesa	16	37	64	0.11	22.25	128.71	14	0.06	0.70	0
Isabela	12	16	28	0.15	23.42	75.65	4670	1.71	0.80	1800
Marchena	8	14	20	0.18	12.25	108.00	130	0.34	0.70	0
Pinta	21	76	134	0.08	105.08	128.32	60	0.65	0.80	0
San Cristobal	21	93	234	0.12	109.08	141.25	557	0.73	4.00	6000
Santa Cruz	23	76	215	0.12	126.33	68.93	986	0.86	2.30	12000
Santiago	24	69	167	0.10	107.83	72.25	572	0.91	1.40	0

3.2 Model Development

In developing our model, we make use of a qualitative representation of the pollination data (considering a simplified case). Therefore, for each of the islands, we convert the data to a binary matrix a_{ij} representing presence ($a_{ij} = 1$) or absence ($a_{ij} = 0$) of interaction between a plant and its pollinator (Table 3.1).

In this section, we will consider the dynamical model for changes in species densities, the adaptive rewiring rules (stochasticity in the model), the numerical simulations and how we assigned the model parameters, the null models with no mechanism of interest and finally, we will give and discuss the obtained results.

3.2.1 Dynamical description

To predict the observed structures of nestedness and modularity for each of the island community, we use a network approach in representing the species interaction in the community, where the nodes are the species and links are the mutualistic interactions (Albert and Barabási, 2002). To capture the dynamical changes in the densities of the species, we employed a Lotka-Voltera model of mutualism which incorporates inter and intra specific competition; Holling's type I (linear) and type II (non-linear) functional responses; and an adaptive behaviour of switching interaction partner (Holling (1959); May (1981); Bascompte *et al.* (2006); Holland *et al.* (2006); Okuyama and Holland

(2008); Bastolla *et al.* (2009); Zhang *et al.* (2011); Suweis *et al.* (2013)). Majorly, we adopted two algorithms of interaction switching, that is optimization and elimination algorithms.

We represent the mutualistic networks as bipartite graphs, where plants are one partite P and animals or pollinators the other partite A . Species in the same partite are in competition with each other while they interact mutualistically with species in the other partite (Bastolla et al. (2009); Suweis et al (2013)). Suppose m is the number of plants species in partite P and n the number of animals species in partite A, the population dynamics of these mutualistic species and their interactions can be represented through Hollings type I, linear functional response as:

$$\frac{dP_i}{dt} = r_i P_i - \sum_{j=1}^{n} \alpha_{ij}^{(P)} P_j P_i + \sum_{j=1}^{m} a_{ij} v_{ij} \beta_{ij}^{(P)} A_j P_i \,, \tag{3.2.1}$$

$$\frac{dA_j}{dt} = r_j A_j - \sum_{i=1}^{m} \alpha_{ij}^{(A)} A_i A_j + \sum_{i=1}^{n} a_{ij} v_{ij} \beta_{ij}^{(A)} P_i A_j \,, \tag{3.2.2}$$

where P_i and A_j are the population densities of plant i and animal j, superscripts (P) and (A) represent plants and animals respectively. From the right-hand side of both equations (3.2.1) and (3.2.2), the first term depicts the population growth where r_i and r_j are the per capita growth rates of species in the absence of competition and mutualism; the second term of the equations modifies the species population by subtraction of the competition term $\alpha_{ij}^{(P)}$ and $\alpha_{ij}^{(A)}$, whose diagonal elements ($\alpha_{ii}^{(P)}$ and $\alpha_{jj}^{(A)}$) are the density-dependent coefficients of plants and animals respectively. Adding the mutualistic effect, which is a linear functional response, the third term of the equations represent the saturation to the benefit of a species by the densities of its partner species, where a_{ij}, is a binary interaction matrix which denotes whether a plant i interacts with an animal j ($a_{ij} = 1$) or not ($a_{ij} = 0$); the preference matrix v_{ij} depicts the probability of an interaction whenever an animal j meets a plant i; the mutualistic interaction strength (benefit) matrix $\beta_{ij}^{(P)}$ and $\beta_{ij}^{(A)}$ describe the benefits obtained by a plant i from an animal j and by an animal j from a plant i respectively.

Extending the above model to the one that incorporates the handling time h, we now have the Holling's type II, non-linear functional response given by:

$$\frac{dP_i}{dt} = r_i P_i - \sum_{j=1}^{n} \alpha_{ij}^{(P)} P_j P_i + \frac{\sum_{j=1}^{m} a_{ij} v_{ij} \beta_{ij}^{(P)} A_j P_i}{1 + h \sum_{k=1}^{m} a_{ik} v_{ik} A_k} \,, \tag{3.2.3}$$

$$\frac{dA_j}{dt} = r_j A_j - \sum_{i=1}^{m} \alpha_{ij}^{(A)} A_i A_j + \frac{\sum_{i=1}^{n} a_{ij} v_{ij} \beta_{ij}^{(A)} P_i A_j}{1 + h \sum_{k=1}^{n} a_{jk} v_{jk} P_k} \,, \tag{3.2.4}$$

where the parameters are defined in the same way as we did above. The Holling's non-linear functional responses which is the third term in the right hand side of equations (3.2.3) and (3.2.4) represent the saturation to the benefit of a species by the densities of the interacting species (Holland *et al.* (2006); Okuyama and Holland (2008); Bastolla *et al.* (2009); Zhang *et al.* (2011); Suweis *et al.* (2013)).
Note that equations (3.2.3 and 3.2.4) is equal to equations (3.2.1 and 3.2.2) when h=0.

With these models above, we have described changes in species densities. Next we now incorporate the behavioural adaptation of species into the model, that is, adaptive interaction switch because species have being observed to switch their interaction partners in response to resource availability and perturbations from the environment. These switching or rewiring adds stochasticity to the models.

3.2.2 Adaptive Rewiring Algorithms

We implement the adaptive behaviour of switching interaction partners using two algorithms namely: elimination and optimization algorithms. Elimination algorithm uses a principle of eliminating unfit species that contributes least to a species fitness, by switching its interactions and then randomly trying out a new species (Zhang *et al.*, 2011) while the optimization algorithm employs a strategy of increasing species richness at equilibrium through the interchange of interaction strengths between species (Suweis *et al.*, 2013).

In details, the elimination algorithm is implemented in this way: at every time step, suppose a pollinator species j is randomly selected, then it terminates an interaction with a plant species i that contributes least to its fitness (that is the least gain in the per capita population growth rate), and then a randomly chosen plant species k that has no interaction with the pollinator species j is assigned an interaction, keeping the number of interactions constant and updating the preference value for the new assigning of interaction. Same procedure holds if a plant species is initially randomly selected. This process of trying out new species ensures a behavioural adaptation for the species, because they would tend to interact with the species that give them more gain and discard less beneficial ones.

For the optimization algorithm, a randomly chosen species i with an existing link to other species potentially switches its interaction and interchanges benefits with the other species to maximize its abundance at equilibrium. The switch is accepted if and only if, it does not lead to a decrease in the population abundance of the species i, otherwise the switch would correspond to just interchange of links. In this algorithm, increases in a species abundance leads

to an increment in the population of the community.

With these constructions, our models are divided into four:

- Hollings type I, Elimination algorithm (IE);
- Hollings type II, Elimination algorithm (IIE);
- Hollings type I, Optimization algorithm (IO);
- Hollings type II, Optimization algorithm (IIO).

From here, we will now perform numerical simulations using the above four AIS models to predict network structures.

3.2.3 Numerical Simulations and Parametrizations

We conducted the numerical simulations, solving the ordinary differential equations with Runge-Kutta order 4 method with a step size of 0.01, using *deSolve* package in R version 3.2.2 (R Development Core Team, http://www.r-project.org). At every time step, the switching algorithms are implemented and network structures are computed (see Appendix A). To capture the observed levels of the network structures, we assigned the input parameters of the models from various distributions such as uniform distribution - U(a,b), normal distribution - $\mathrm{N}(\mu, \sigma)$ and log-normal distribution - $\mathrm{LN}(\mu, \sigma)$, where a and b are minimum and maximum values, μ is the mean and σ is the standard deviation.

In the numerical simulations of our four AIS models, the predictions of network structures is obtained by incorporating the interaction switches and solving the ordinary differential equations. Initial parameters (input) is assigned as follows: we randomly generated the binary interaction matrix a_{ij} with the numbers of plants, animals and interactions obtained from the empirical data; initial population size densities is assigned to follow $\mathrm{U}(0, 1)$; intrinsic growth rates $r_i, r_j \sim \mathrm{LN}(0,0.1)$; the preference matrix $v_{ij} \sim \mathrm{U}(0,1)$; the mutualistic interaction strength (benefit) matrices $\beta_{ij}^{(P)}, \beta_{ij}^{(A)} \sim |\mathrm{N}(0,0.01)|$; the competition matrices $\alpha_{ij}^{(P)}, \alpha_{ij}^{(A)} \sim |\mathrm{N}(0,0.01)|$ and we assigned the coefficient of the density dependences $\alpha_{ii}^{(P)} = 1, \alpha_{jj}^{(A)} = 1$. In the normal distribution, the choice of standard deviation of 0.01 for the inter-species interaction is to ensure co-existence among the species, especially for the models with type I functional response (IE and IO). The handling time for all species were assigned the same, that is for models with type I and type II functional responses, h=0 and h=0.1 respectively, using equations (3.2.3) and (3.2.4).

To identify groups of interacting species in each island, we computed for modularity. This provides insight into the structure of the mutualistic communities. (Olesen *et al.*, 2007). For each island network, modularity partitioned the species into different compartments or modules with species in the same module interacting more often than expected. In achieving this partitioning, we use simulated annealing in NETCARTO software to maximize the computed modularity value for each island network (Guimera & Nunes Amaral 2005; Guimera & Amaral 2005). Also to account for how specialist species interacted with a subset of generalist species, we used NODF2 nested function from *bipartite* package in R (Almeida-Neto et al. 2008). NODF2 sorts the interaction matrix before finding its nestedness.

Generally, we kept the number of interactions constant. In one simulation, we simulated over 10 000 time steps, switching interactions in each time step. Once a network structure has converged to its value, the model prediction for this network structure is the average of the last 100 matrices. Because the model is stochastic, we repeated the simulations for 100 times, capturing different model predictions for each island network (see Table 3.2 and Table 3.4 for model predictions).

3.2.4 Null models

To know the significance of the nestedness and modularity predictions from the rewiring models, a comparison is needed with null models that have no special mechanism of interest. To do this, we made use of two null models, Null4 and Null5 from *bipartite* package in R. Null4 shuffles a binary matrix maintaining it dimensions and keeping the number of interactions constant, but the marginal sums are not kept constant (Gotelli *et al.* (1996); Fortuna and Bascompte (2006)) while Null5 is based on a probability matrix in the generation of the binary networks with a desired number of interaction, where marginal sums and connectance are not constant (Vázquez *et al.*, 2009).

Using these two null models, we generated 1000 random matrices for each Island network and then computed the nestedness and modularity values of the matrices using NODF2 and NETCARTO as explained in the previous subsection. The null model predictions are the averages over the 1000 matrices for each of the network structure. (Tables 3.2 and 3.4).

Therefore, to access the predictions from the AIS models and the null models, we use Reduced Major Axis (RMA) regression from *lmodel2* package in R. The RMA helps us to know how well the models are predicting the observed structures.

Having described the AIS models and the null models, we now give the results

we obtained from their predictions of the network structures.

3.2.5 Results

Here we give the results of the numerical simulations of the Island data. Results from the predictions of nestedness and modularity will be highlighted including results on the persistence of the species.

3.2.5.1 Nestedness

Nested structures were predicted from the AIS models. We now explain how different initial random networks converged to the observed nested values. Thereafter, we give the percentages of the explained variance in the observed networks from the AIS model predictions and null model predictions.

3.2.5.1.1 Different initial values convergence: The predicted nestedness of the pollination networks for some of our AIS models gradually converged to the observed level of the empirical data (Fig. 3.2a, but see Fig. 3.3a). Consequently, all model predictions were significantly different from their observed values (Table 3.2). When simulations were executed from different initial network structures and model parameters, the predictions attained equilibrium. This suggests that our rewiring models are robust. We found from the elimination algorithm, that only three island networks predictions converged to their observed nested values (that is for Española, Floreana and Marchena island networks. See Fig. 3.4 for an example). The remaining seven island networks were underestimated in their predicted values (see Fig. 3.5 for an example), which could potentially be as a result of the low connectance values (< 0.18) of the seven islands (Table 3.1). Furthermore, the optimization algorithm gave higher predictions of the observed networks (see Fig. 3.6 for an example) except for Floreana, Isabela and Marchena islands networks (Table 3.2).

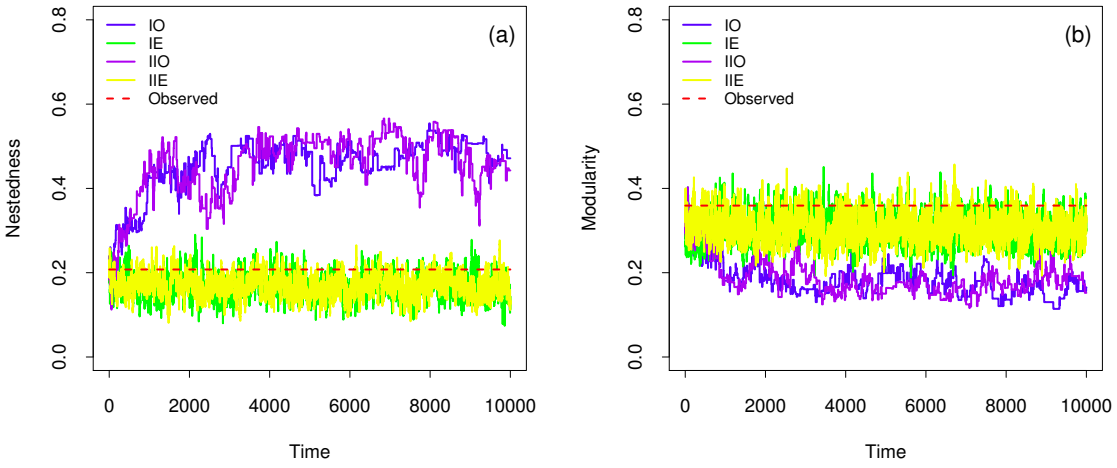


Figure 3.2: Nestedness and Modularity dynamics of the four rewiring algorithms (Adaptive Interaction Switch models). Model simulations are done using Española island data. Red broken line is the observed value of each structure. Elimination models are IE (type I functional response) and IIE (type II functional response); optimization models are IO (type I functional response) and IIO (type II functional response).

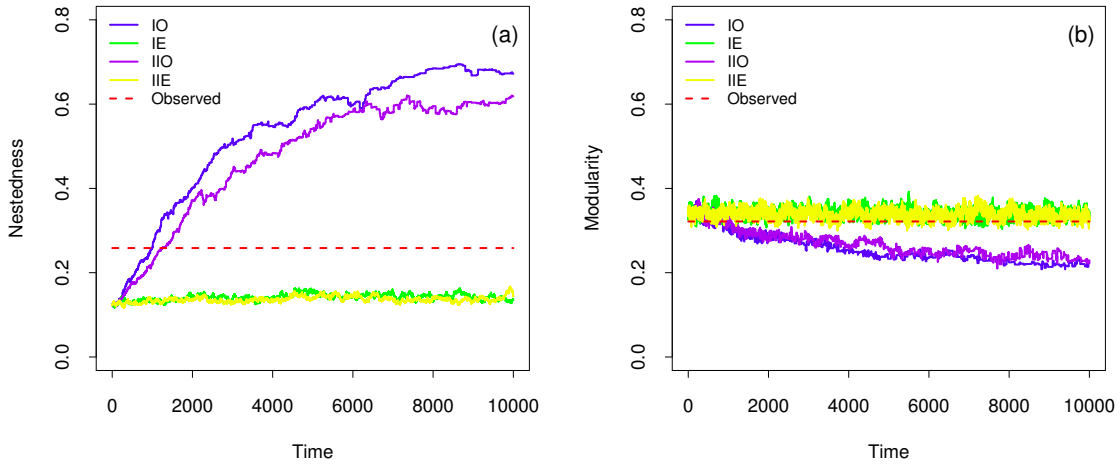


Figure 3.3: Nestedness and Modularity dynamics of the four rewiring algorithms (Adaptive Interaction Switch models). Model simulations are done using Santacruz island data. Red broken line is the observed value of each structure. Elimination models are IE (type I functional response) and IIE (type II functional response); optimization models are IO (type I functional response) and IIO (type II functional response).

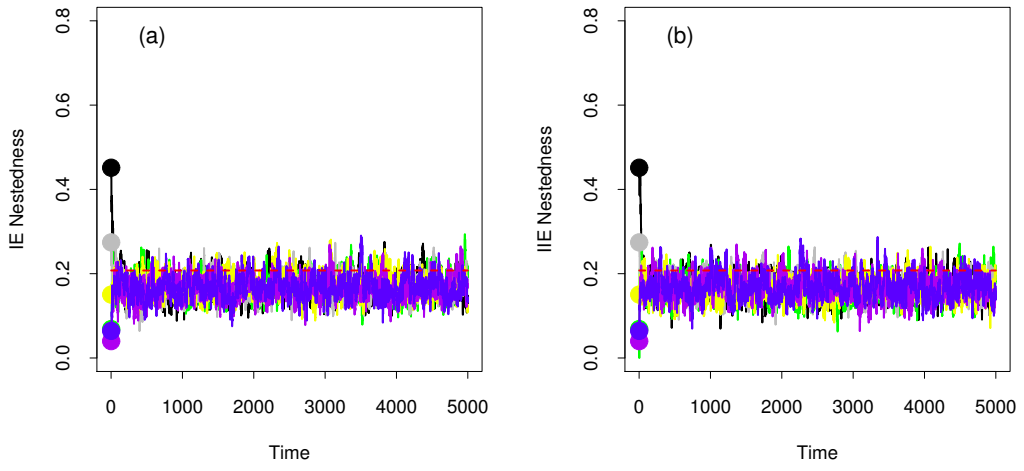


Figure 3.4: Nestedness dynamics of Española Island using the elimination algorithm models (IE and IIE). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of nestedness. Model prediction fitted the observed level.

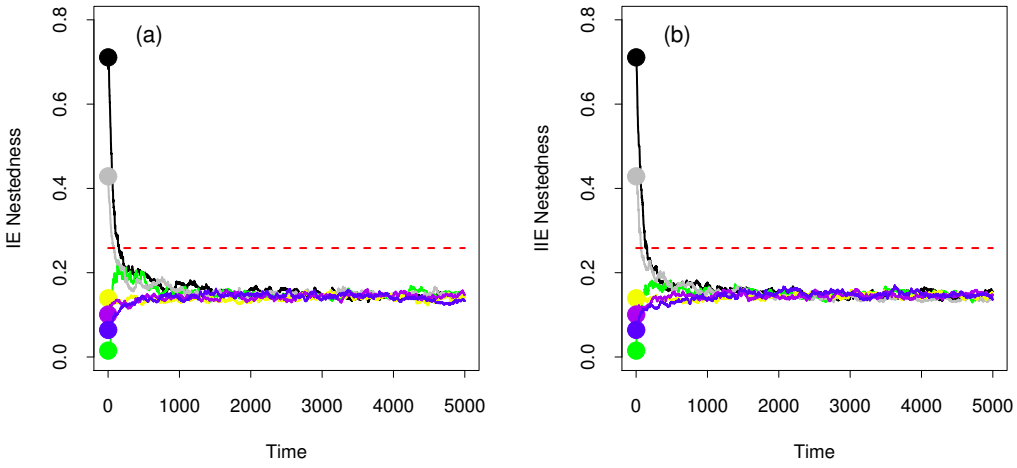


Figure 3.5: Nestedness dynamics of Santacruz Island using elimination algorithm models (IE and IIE). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of nestedness. Model prediction underestimated the observed level.

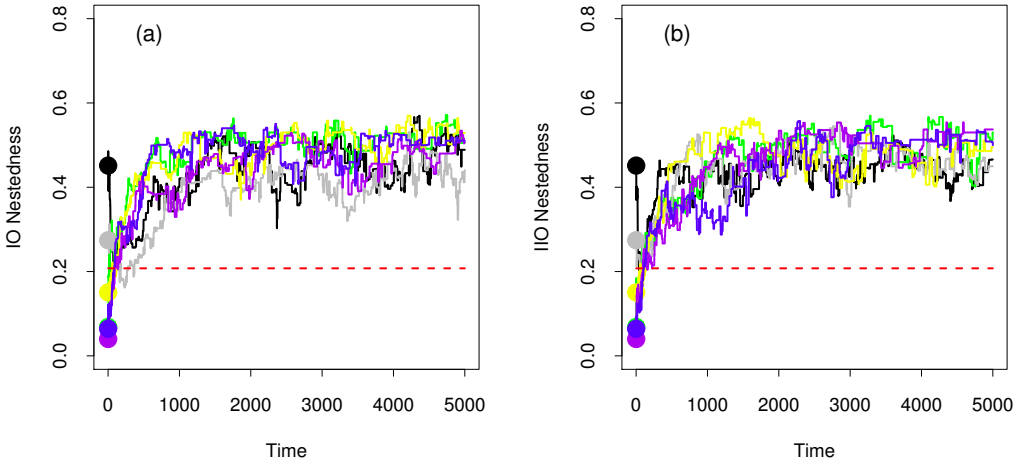


Figure 3.6: Nestedness dynamics of Española Island using optimization algorithm models (IO and IIO). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of nestedness. Model prediction overestimated the observed level.

3.2.5.1.2 Nestedness predictions: Accessing the predictions from each of the four AIS models with the observed nestedness values, some level of variations in the observed were explained. According to the Reduced Major Axis (RMA) regression, the elimination models explained almost 40% of the variation in the observed nestedness, that is IE (r^2= 0.39, RMA slope = 1.3, p = 0.054) and IIE (r^2 = 0.38, RMA slope = 1.38, p = 0.056) at a 10% significance level (Fig. 3.7). In contrast, the optimization models captured less than 1% of the explained variance of no significance (Table 3.3). The optimization algorithm overestimated nestedness values and thus does not fit the observed data. Even the null models accounted for some level of variations that is Null4 (r^2= 0.31, RMA slope = 0.91, p = 0.09) and Null5 (r^2= 0.28, RMA slope = 1.71, p = 0.12) (Fig. 3.7).

Table 3.2: Nestedness prediction values for our four Adaptive Interaction Switch (AIS) models and two null models, showing the 95% *confidence interval* $\left(\bar{X} \pm Z_{\alpha/2}\frac{\sigma_{\bar{X}}}{\sqrt{n}}\right)$, *where* $\bar{X}$ *and* $\sigma_{\bar{X}}$ *are the mean and standard deviation of the model predictions respectively;* $Z_{\alpha/2} = 1.96$ *is the* 95% *confidence value; number of samples n is 100 for our AIS models and 1000 for the null models. Elimination models are IE (type I functional response) and IIE (type II functional response); optimization models are IO (type I functional response) and IIO (type II functional response), null models are NULL4 and NULL5. All predictions are significantly different from the observed values* ($p < 0.0001$).

Islands	Observed	IO	IE	IIO	IIE	Null4	Null5
Espanola	0.2076	0.4408 ± 0.0153	0.1690 ± 0.0030	0.4455 ± 0.0127	0.1673 ± 0.0027	0.1863 ± 0.0009	0.1621 ± 0.0020
Fernandina	0.1938	0.5193 ± 0.0183	0.1254 ± 0.0013	0.5232 ± 0.0138	0.1250 ± 0.0013	0.1205 ± 0.0002	0.1220 ± 0.0006
Floreana	0.2705	0.2019 ± 0.0097	0.1545 ± 0.0021	0.1694 ± 0.0089	0.1527 ± 0.0019	0.1791 ± 0.0013	0.1335 ± 0.0037
Genovesa	0.1737	0.3405 ± 0.0170	0.1007 ± 0.0013	0.3638 ± 0.0120	0.1045 ± 0.0014	0.111 ± 0.0003	0.1036 ± 0.0008
Isabela	0.2428	0.2879 ± 0.0099	0.1260 ± 0.0023	0.2856 ± 0.0097	0.1260 ± 0.0023	0.1528 ± 0.0008	0.1255 ± 0.0016
Marchena	0.1863	0.2264 ± 0.0093	0.1315 ± 0.0023	0.2254 ± 0.0106	0.1310 ± 0.0024	0.1283 ± 0.0009	0.1196 ± 0.0018
Pinta	0.1791	0.3773 ± 0.0189	0.0844 ± 0.0008	0.4004 ± 0.0147	0.0840 ± 0.0009	0.0872 ± 0.0002	0.0841 ± 0.0004
San Cristobal	0.2468	0.5890 ± 0.0202	0.1385 ± 0.0010	0.5577 ± 0.0185	0.1392 ± 0.0011	0.1365 ± 0.0001	0.1295 ± 0.0004
Santa Cruz	0.2585	0.6007 ± 0.0225	0.1448 ± 0.0011	0.5922 ± 0.0133	0.1443 ± 0.0011	0.1233 ± 0.0002	0.1334 ± 0.0005
Santiago	0.2056	0.5275 ± 0.0208	0.1141 ± 0.0012	0.5485 ± 0.0126	0.1153 ± 0.0011	0.1050 ± 0.0002	0.1090 ± 0.0004

Table 3.3: Nestedness linear regression results showing the comparison between the observed values and the model predictions, their R squared values, slope and p-values, using Reduced Major Axis (RMA) linear regression. R squared is the coefficient of determination. Elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5.

Models	R squared	RMA slope	p-values
Observed vs IO	0.0102	0.0337	0.7813
Observed vs IE	0.3889	1.3031	0.0540
Observed vs IIO	0.0002	-0.0044	0.9707
Observed vs IIE	0.3840	1.3769	0.0560
Observed vs Null4	0.3143	0.9127	0.0919
Observed vs Null5	0.2798	1.7146	0.1159

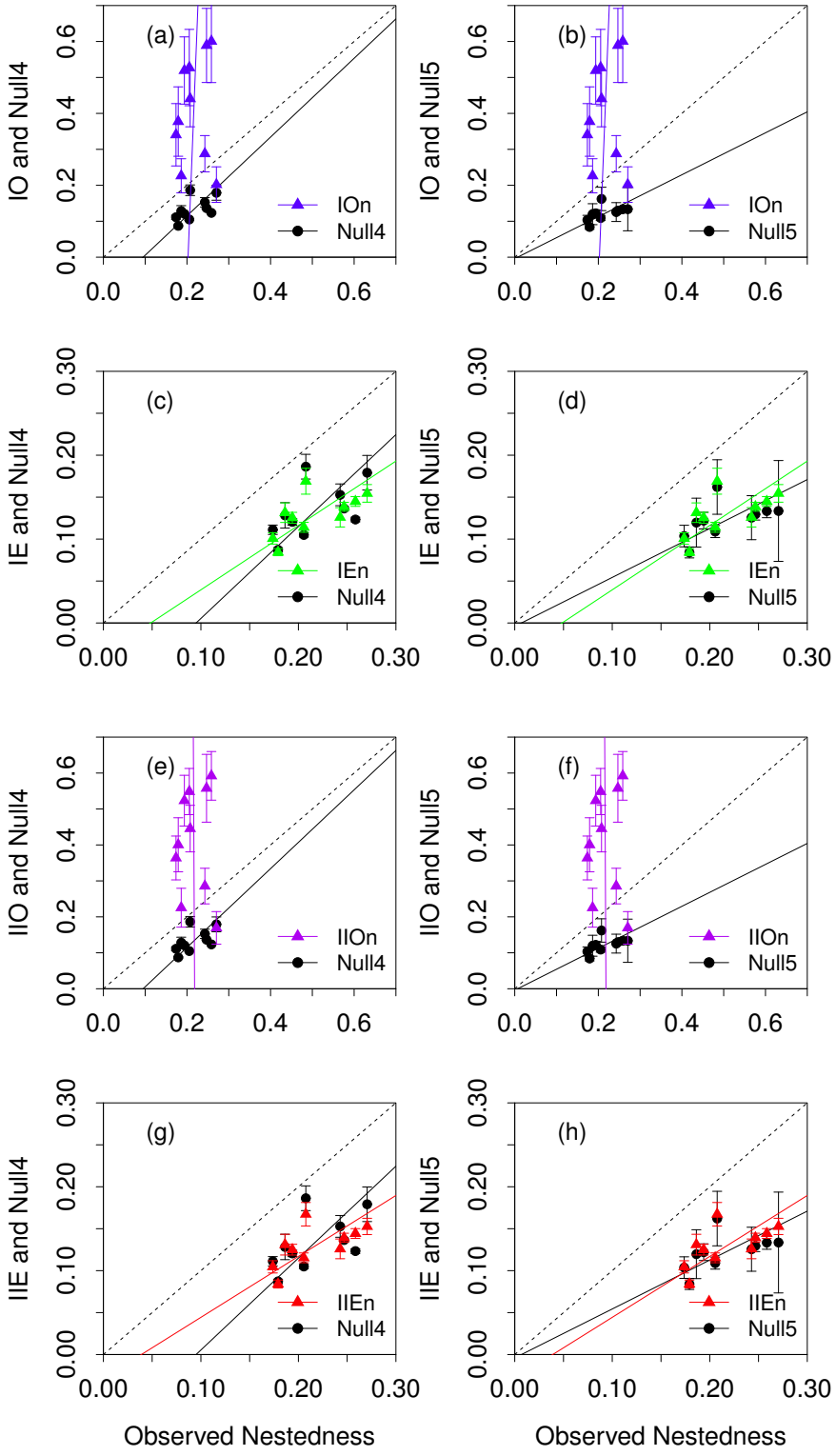


Figure 3.7: Nestedness predictions comparison between each of our four Adaptive Interaction Switch (AIS) models and the two null models against the observed values. Each dot represents an island network. Elimination models are IE (linear functional response) and IIE (nonlinear functional response); optimization models are IO (linear functional response) and IIO (nonlinear functional response); null models are Null4 and Null5. The broken line is for $y = x$.

3.2.5.2 Modularity

Modular structures were predicted from the AIS models. We now explain how different initial random networks converged to the observed modular values. Thereafter, we give the percentages of the explained variance in the observed networks from the AIS and null model predictions.

3.2.5.2.1 Different initial values convergence: The predicted level of modularity of the pollination networks gradually converged to the observed level of the empirical data (Fig. 3.2b and Fig. 3.3b). Moreover, all model predictions were significantly different from their observed values except for the IO model prediction of Floreana island network (Table 3.4). When simulations were executed from different initial network structures and model parameters, the predictions attained equilibrium. This suggests that our rewiring models are robust. In addition, predictions from the elimination algorithm, converged to their observed modularity values for all island networks (see Fig. 3.8 for an example). Even with the low connectance values of some of the islands, there was no underestimation of any island network unlike the case of the nestedness predictions. In contrast, the optimization models underestimated the observed modularity (see Fig. 3.9 for an example) of the island networks except for Floreana, Isabela and Marchena islands unlike its nestedness predictions.

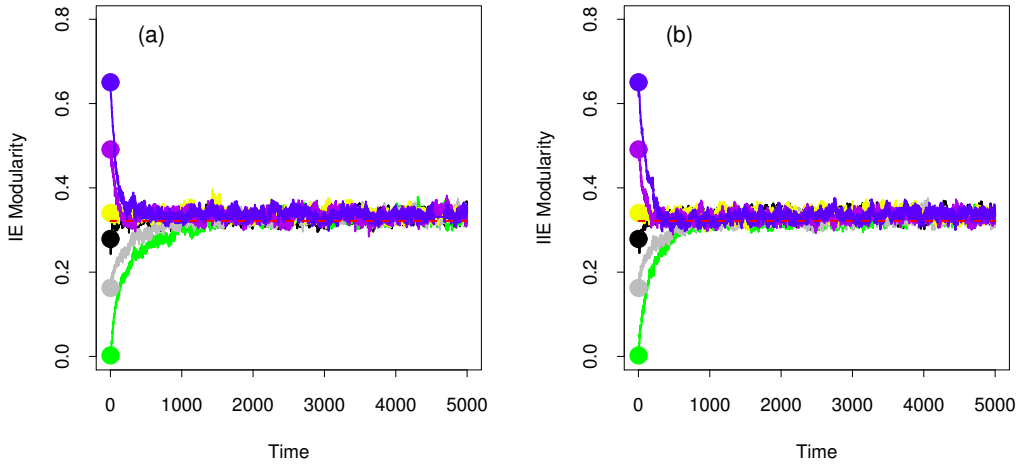


Figure 3.8: Modularity dynamics of Santa Cruz Island using Elimination algorithm models (IE and IIE). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of modularity. Model prediction fitted the observed level.

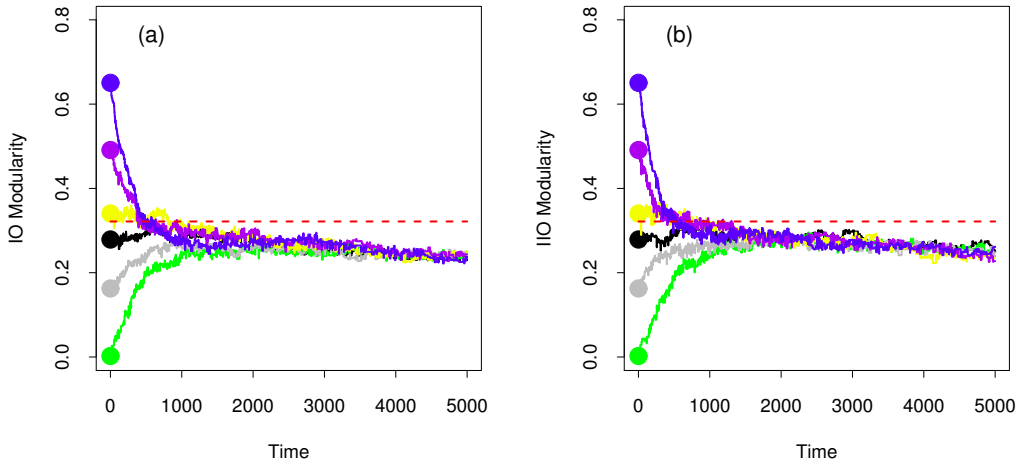


Figure 3.9: Modularity dynamics of Santa Cruz Island using Optimization algorithm models (IO and IIO). (a) is for type I functional response and (b) is for type II functional response. Simulations started with different initial network structures that is black, grey, green, yellow, blue and purple lines. Red broken line is the observed level of modularity. Model prediction underestimated the observed level.

3.2.5.2.2 Modularity predictions: Accessing the predictions from each of the four AIS models with the observed modularity values, more level of variations in the observed were explained. According to the RMA regression, the elimination models explained more than 50% of the variation in the observed modularity, that is IE (r^2= 0.521, RMA slope = 1.09, p = 0.018) and IIE (r^2= 0.523, RMA slope = 1.08, p = 0.018) at a 5% significance level (Fig 3.10). From the optimization models, some level of variation was explained by IO (r^2= 0.41, RMA slope = 1.1 p = 0.046) and a lesser variation was explained by IIO (r^2 = 0.19, RMA slope = 0.19, p = 0.21). Surprisingly, the null models accounted for a higher level of variations that is Null4 (r^2= 0.62, RMA slope = 1.03, p = 0.007) and Null5 (r^2= 0.59, RMA slope = 1.1, p = 0.009) (Table 3.5; Fig. 3.10).

Table 3.4: Modularity prediction values for our four Adaptive Interaction Switch (AIS) models and two null models, showing the 95% confidence interval $\left(\bar{X} \pm Z_{\alpha/2}\frac{\sigma_{\bar{X}}}{\sqrt{n}}\right)$*, where* $\bar{X}$ *and* $\sigma_{\bar{X}}$ *are the mean and standard deviation of the model predictions respectively;* $Z_{\alpha/2} = 1.96$ *is the 95% confidence value; number of samples n is 100 for our AIS models and 1000 for the null models. Elimination models are IE (type I functional response) and IIE (type II functional response); optimization models are IO (type I functional response) and IIO (type II functional response), null models are NULL4 and NULL5. All predictions are significantly different from the observed values (*$p < 0.0001$*), except for the underlined prediction (*$p = 0.845$*).*

Islands	Observed	IO	IE	IIO	IIE	Null4	Null5
Espanola	0.3592	0.2008 ± 0.0073	0.3054 ± 0.0032	0.2029 ± 0.0059	0.3050 ± 0.0030	0.3048 ± 0.0020	0.3231 ± 0.0023
Fernandina	0.3884	0.2921 ± 0.0053	0.4046 ± 0.0018	0.2934 ± 0.0043	0.4042 ± 0.0016	0.4041 ± 0.0009	0.4180 ± 0.0011
Floreana	0.2851	**0.2863 ± 0.0115**	0.3715 ± 0.0043	0.3376 ± 0.0143	0.3727 ± 0.0041	0.3407 ± 0.0031	0.3730 ± 0.0040
Genovesa	0.3867	0.3176 ± 0.0089	0.4345 ± 0.0026	0.3075 ± 0.0061	0.4353 ± 0.0026	0.4329 ± 0.0014	0.4459 ± 0.0015
Isabela	0.2804	0.2126 ± 0.0099	0.3418 ± 0.0033	0.2159 ± 0.0107	0.3413 ± 0.0034	0.3305 ± 0.0022	0.3501 ± 0.0025
Marchena	0.3102	0.2730 ± 0.0086	0.3699 ± 0.0038	0.2736 ± 0.0102	0.3666 ± 0.0032	0.3744 ± 0.0027	0.3836 ± 0.0030
Pinta	0.4514	0.3528 ± 0.0084	0.4786 ± 0.0018	0.3479 ± 0.0062	0.4789 ± 0.0018	0.4786 ± 0.0009	0.4906 ± 0.0010
San Cristobal	0.3523	0.2671 ± 0.0046	0.3660 ± 0.0012	0.2730 ± 0.0044	0.3658 ± 0.0014	0.3666 ± 0.0007	0.3837 ± 0.0007
Santa Cruz	0.3216	0.2487 ± 0.0046	0.3344 ± 0.0014	0.2471 ± 0.0033	0.3356 ± 0.0015	0.3390 ± 0.0007	0.3511 ± 0.0008
Santiago	0.3333	0.2804 ± 0.0050	0.3747 ± 0.0017	0.2739 ± 0.0033	0.3750 ± 0.0017	0.3783 ± 0.0009	0.3888 ± 0.0009

Table 3.5: Modularity linear regression results showing the comparison between the observed values and the model predictions, their R squared values, slope and p-values, using Reduced Major Axis (RMA) linear regression. R squared is the coefficient of determination. Elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5.

Models	R squared	RMA slope	p-values
Observed vs IO	0.4103	1.0986	0.0460
Observed vs IE	0.5213	1.0856	0.0184
Observed vs IIO	0.1857	0.9135	0.2137
Observed vs IIE	0.5225	1.0795	0.0182
Observed vs Null4	0.6212	1.0267	0.0068
Observed vs Null5	0.5929	1.1087	0.0092

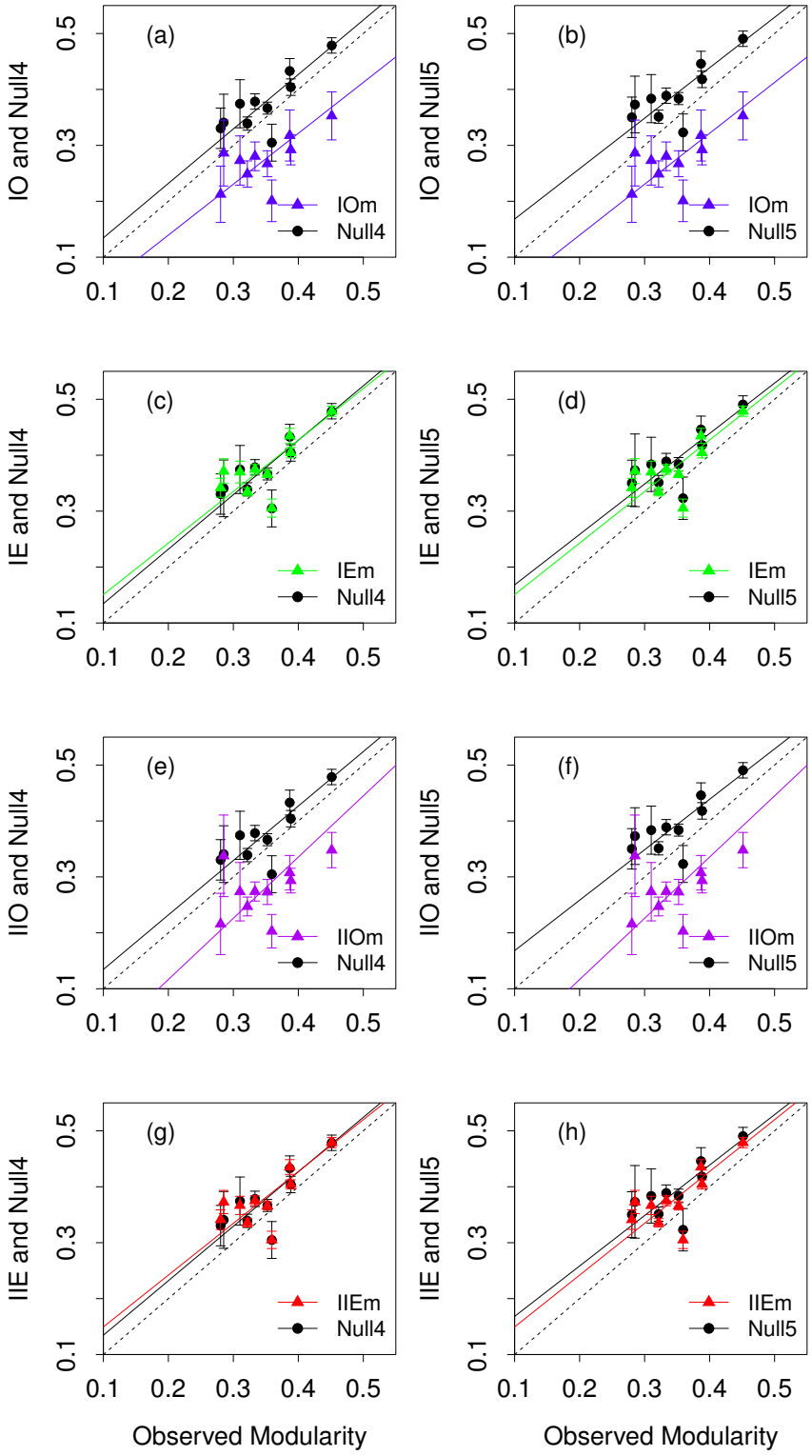


Figure 3.10: Modularity predictions comparison between each of our four Adaptive Interaction Switch (AIS) models and the two null models against the observed values. Each dot represents an island network. Elimination models are IE (linear functional response) and IIE (nonlinear functional response); optimization models are IO (linear functional response) and IIO (nonlinear functional response); null models are Null4 and Null5. The broken line is for $y = x$.

3.2.5.3 Persistence

In general, the equilibrium densities of all species persisted with no species going extinct (see Fig. 3.11 for an example). This gives a stability measure of our AIS model predictions.

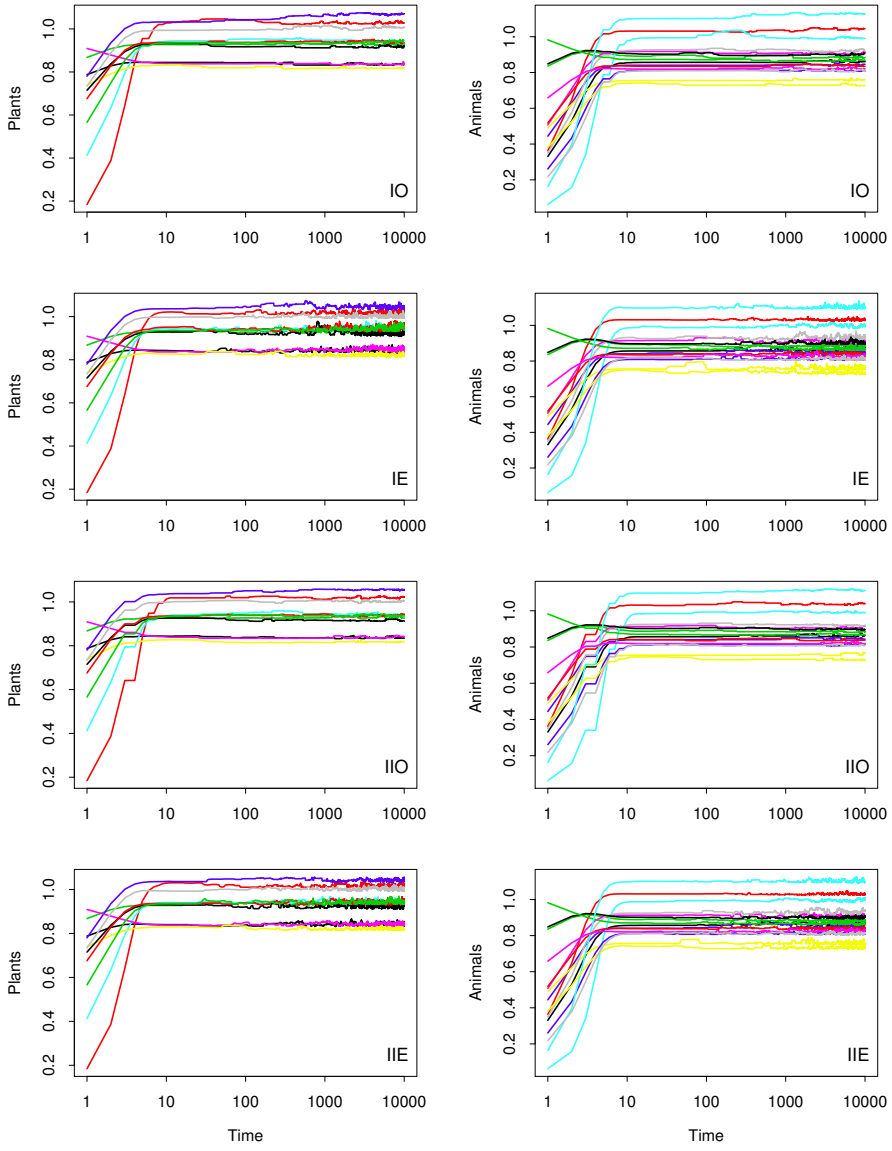


Figure 3.11: Log(x) plot of species population densities of the Adaptive Interaction Switch models using Española Island data. The left column figures are for plant species while the right column figures are for animal species. The first and third row figures are for optimization models IO and IIO respectively; the second and fourth row figures are for elimination models IE and IIE respectively.

3.2.6 Discussion

Here, we have used a dynamical model to represent biotic interactions of mutualistic species and predicted the emergence of structures using 10 pollination data from Galápagos islands. This we achieved by using a Lotka-Voltera model of mutualism, and incorporated species competition, functional responses (Holling type I and type II) and an adaptive interaction switch (AIS) (Holling (1959); May (1981); Okuyama and Holland (2008); Bastolla *et al.* (2009); Zhang *et al.* (2011); Suweis *et al.* (2013)). Behavioural adaptation of switching have being found in pollination networks (Basilio *et al.* (2006); Petanidou *et al.* (2008)), not only allowing a static community to become dynamic through the interchange of links, but effectively predicting the emergence of network structures. In consistent with some previous studies (Zhang *et al.* (2011); Suweis *et al.* (2013); Nuwagaba *et al.* (2015)), our AIS models predicted nestedness and modularity from different initial network structure and model parameters. The dynamics of the predicted structures converged to the observed level. This suggests the importance of AIS in predicting network structures. Moreover, there was no extinction of species as all the species persisted throughout the switching. Although a recent study (Gilljam *et al.* (2015)) showed that species rewiring can intensify the effect of species loss in foodwebs, creating pressures on rewired species. Nevertheless, adaptive rewiring have a beneficial effect in mutualistic networks (Zhang *et al.* (2011)). Hence AIS is a key eco-evolutionary process in structural emergence (Valdovinos *et al.* (2010)).

Observed network connectance correlated positively with the observed nested values and negatively with the observed values of modularity (Fig. 3.12). As a result of this correlations, predicted networks with lower connectance tend to have low nested predictions because they were underestimated in their nested predictions as in the case of the elimination algorithm which explained almost 40% variation in the observed nestedness. More than half of the variation in the observed modularity was explained by the elimination algorithm which could potentially be as a result of the low connectance values of the networks as noted by Olesen *et al.* (2007), suggesting that low connectance in network increases modularity. Although optimization model overestimated observed nestedness, providing unrealistic estimates and hence fitting poorly to the observed data, it accounted for some variations in the observed modularity, suggesting that maximizing species richness have more effect in predicting the modularity of a network than the nestedness.

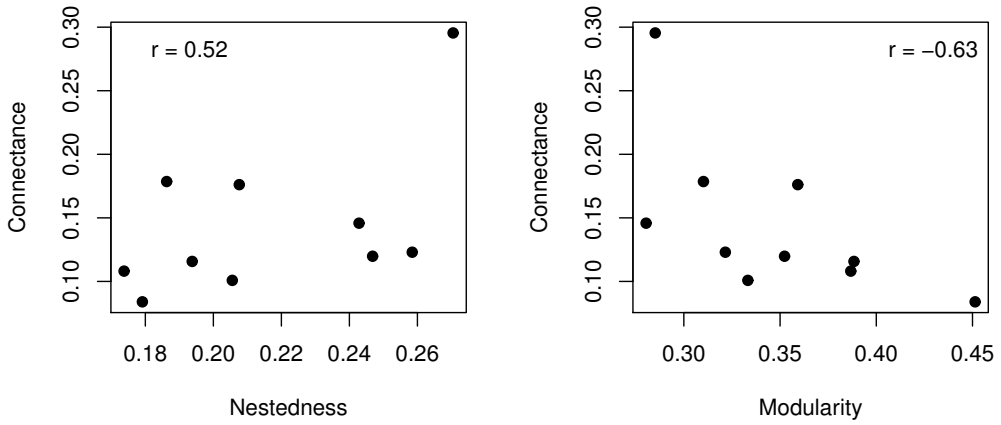


Figure 3.12: Correlation of the observed connectance with observed nestedness and modularity, r is the correlation coefficient.

In conclusion, we have predicted network structures using AIS models and Null models, we now move on to the next Chapter to add environmental variables together with the models as predictors of network structures.

Chapter 4

Statistical Analysis of the Environmental Variables

Environmental variables can have effects in predicting network structures, as such, it is in our interest to investigate these effects. These variables include geographic factors (e.g Island size, isolation, age and elevation) and anthropogenic factors (e.g sampling effort and human population size). In addition to the AIS model predictions, we now see how these variables are contributing to the explained variation in the observed network structures, using a Generalized linear model (GLM). In this Chapter, we start by describing the data on the variables, thereafter we state the results we obtained and discuss it.

4.1 Data

With respect to the Galápagos island community, we compiled data on the environmental variables from online and published sources. Using google earth (7.1) software, we measured the average nearest distances for each island to the other Islands. This nearest distance (in km) represents the isolation of the Islands. Informations on Island size (that is, area in km^2), maximum elevation (in km) and human population size was obtained from Conservancy (2016). We obtained the Island age which was over minimum and maximum ranges from Poulakakis *et al.* (2012) and Conservancy (2016) and then we chose to work with the maximum values in the ranges. Finally, the sampling efforts for each of the islands ranged from 6 - 109 hours (Table 3.1).

Having compiled the data on six environmental variables, we now describe the methods we adopted in analysing the data.

4.2 Method of Analysis

Generally, we used different packages in R to analyse our data on the 6 environmental variables.

Firstly, we started by checking for collinearity among the explanatory variables (AIS and null model predictions and environmental variables) by using VIF function (Variance Inflation Factor) in *fmsb* package in R. To obtain covariants that are independent, any variable contributing to a VIF value greater than 5 is removed. After this checking, we were left with the explanatory variables that were independent of each other (VIF < 5).

Thereafter, we used these remaining explanatory variables to predict the observed network structures. This we achieved by conducting a Generalized Linear Model (GLM), to know the effect of these explanatory variables in explaining the observed values of the network structure.

Furthermore, to investigate the independent contribution of each of explanatory variables to the total explained variance in the observed values of network structures, we partitioned the explained variance obtained from the GLM results using *hier.part* package in R.

From the above analysis, we obtained the following results as we will see in the next section.

4.3 Results

We now give the results from the above analysis on the nestedness predictions and the modularity predictions.

4.3.1 Nestedness

Here we start by stating the VIF results obtained from the explanatory variables and then we give the multiple regression results.

4.3.1.1 Variance Inflation Factor (VIF)

From the variance inflation factor (VIF) tests of all the explanatory variables which include the AIS and null model predictions, island area, nearest distance, age, maximum elevation, sampling effort and human population size, we dropped age and maximum elevation to maintain low VIF values less than 5 (Table 4.1).

Table 4.1: Nestedness Variance Inflation Factor (VIF) showing the VIF values for each combination of the explanatory variables. ND is the nearest distance, SE is the sampling effort, AR is the area, HPS is the human population size (log transformed [$log(x+1)$]), ME is the maximum elevation, AG is the age. To obtain non collinearity among the variables, we make sure to achieve VIF values less than 5, dropping variables with high VIF values. (a) is the combination of all the environmental variables and the model predictions. (b) is the combination of the covariants without age. (c) is the combination of the covariants without age and maximum elevation. Elimination models are IE (type I, linear functional response) and IIE (type II, non-linear functional response); optimization models are IO (type I, linear functional response) and IIO (type II, non-linear functional response); null models are Null4 and Null5. R squared is the coefficient of determination of the Generalized linear model fittings.

(a) Covariant	IO	IE	IIO	IIE	Null4	Null5
ND ~ AR + ME + HPS + AG + SE + model	2.53	3.81	2.53	3.85	3.17	2.85
AR ~ ND + ME + HPS + AG + SE + model	4.96	8.74	5.09	8.56	7.14	5.48
ME ~ ND + AR + HPS + AG + SE + model	4.02	8.15	3.86	7.78	13.16	5.59
HPS ~ ND + AR + ME + AG + SE + model	3.6	3.36	3.93	3.37	3.36	3.42
AG ~ ND + AR + ME + HPS + SE + model	3.78	10.21	3.58	10.21	22.5	7.66
SE ~ ND + AR + ME + HPS + AG + model	5.08	7.48	5.58	7.48	25.47	4.42
Model ~ ND + AR + ME + HPS + AG + SE	4.46	6.48	4.21	6.54	23.73	4.09
R squared of glm(observed ~ model + other variables)	0.97	0.96	0.96	0.96	0.96	0.96

Without age

(b) Covariant	IO	IE	IIO	IIE	Null4	Null5
ND ~ AR + ME + HPS + SE + model	2.05	1.8	2.01	1.79	2.03	1.87
AR ~ ND + ME + HPS + SE + model	4.96	5.95	5.09	5.91	5.18	5.11
ME ~ ND + AR + HPS + SE + model	3.28	3.54	3.28	3.49	3.46	3.34
HPS ~ ND + AR + ME + SE + model	1.75	3.01	1.79	3.04	3.05	2.4
SE ~ ND + AR + ME + HPS + model	4.91	2.98	5.43	2.98	4.15	2.51
Model ~ ND + AR + ME + HPS + SE	3.59	1.93	3.57	1.95	3.21	1.63
R squared of glm(observed ~ model + other variables)	0.85	0.89	0.85	0.89	0.95	0.88

Without age and maximum elevation

(c) Covariant	IO	IE	IIO	IIE	Null4	Null5
ND ~ AR + HPS + SE + model	2.04	1.79	1.99	1.79	2.03	1.87
AR ~ ND + HPS + SE + model	2.2	2.42	2.25	2.45	2.18	2.19
HPS ~ ND + AR + SE + model	1.63	2.6	1.66	2.66	2.69	2.19
SE ~ ND + AR + HPS + model	4.44	2.01	4.7	2.04	3	1.79
Model ~ ND + AR + HPS + SE	3.59	1.79	3.57	1.83	3.04	1.6
R squared of glm(observed ~ model + other variables)	0.79	0.86	0.79	0.85	0.92	0.83

4.3.1.2 Multiple regression

In general, the Generalised Linear Model (GLM) fitting of the observed nested values together with the remaining explanatory variables (model predictions, nearest distance, sampling effort, area and human population sizes) revealed higher explained variance of the observed data. Various combinations explained significantly the observed level of nestedness, from which we obtained the following: for the elimination algorithm, IE + other variables ($R^2 = 0.85$,

$p < 0.001$), IIE + other variables ($R^2 = 0.86$, $p < 0.001$); for optimization algorithm, IO + other variables ($R^2 = 0.79$, $p < 0.01$), IIO + other variables ($R^2 = 0.79$, $p < 0.01$); for the null models, we obtained, Null4 + other variables ($R^2 = 0.92$, $p < 0.0001$), Null5 + other variables ($R^2 = 0.83$, $p < 0.01$) (Table 4.2). These suggest the important role the explanatory variables (environmental factors) are playing in explaining network structures.

Furthermore, we divided the total explained variance among the five explanatory variables. This allowed us to see the independent contribution of each of the covariant to the total explained variance (Fig. 4.1). In this partitioning, individual model contributions are IE = 30.5%, IIE = 30.1%, IO = 1.2%, IIO = 1.04%, Null4 = 30.7%, Null5 = 22% (Table 4.3). Significantly, human population contributed the most to the total explained variance (50% < x < 85%). However, sampling effort, nearest distance and area contributed a lesser percentage to the total explained variance in nestedness (Table 4.3).

Table 4.2: Nestedness multiple regression showing the Generalised linear model (Glm) fitting of the observed values with the model predictions, ND is the nearest distance, SE is the sampling effort, AR is the area, HPS is the human population size (log transformed [$log(x+1)$]). R squared is the coefficient of determination, AIC is the Akaike Information Criterion. Elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5.

Glm	R squared	AIC	p-values
Observed ~ IO + ND + AR + SE + HPS	0.7878	-46.7200	0.0110
Observed ~ IE + ND + AR + SE + HPS	0.8601	-50.3090	0.0002
Observed ~ IIO + ND + AR + SE + HPS	0.7869	-46.7200	0.0114
Observed ~ IIE + ND + AR + SE + HPS	0.8542	-49.9700	0.0003
Observed ~ Null4 + ND + AR + SE + HPS	0.9244	-52.1220	<0.0001
Observed ~ Null5 + ND + AR + SE + HPS	0.8320	-47.8520	0.0014

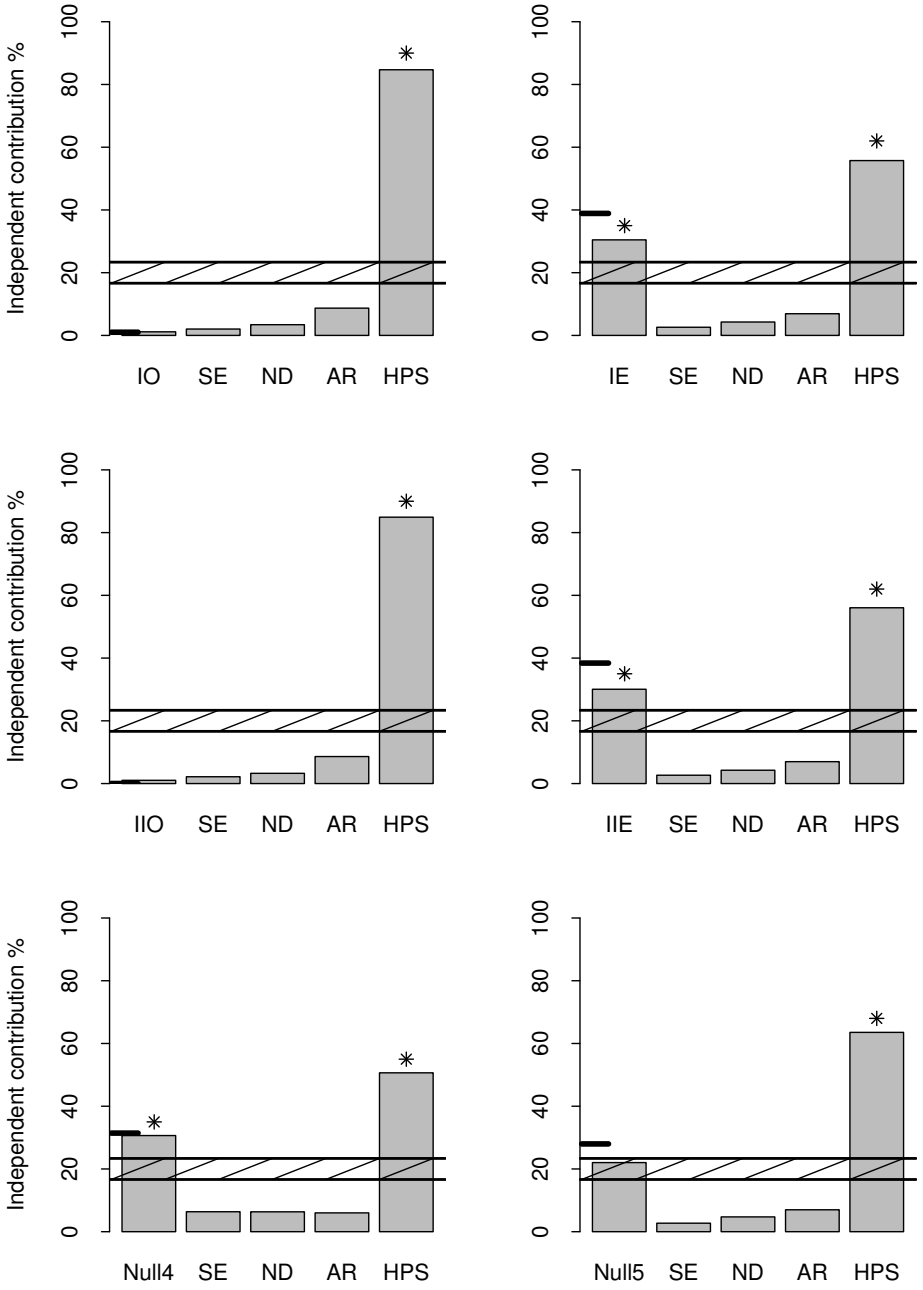


Figure 4.1: Nestedness variance partitioning showing the independent contribution of each of the explanatory variables to the total variance explained for different model combinations. The environmental variables are SE - sampling effort, ND - nearest distance (isolation), AR - area and HPS - human population size (log transformed [log(x+1)]). The band strips shows the 95% confidence interval of the randomly divided variance explained. Bars above this band are important and asterisked bars are significant ($p < 0.1$). Elimination models are IE (linear functional response) and IIE (nonlinear functional response); optimization models are IO (linear functional response) and IIO (nonlinear functional response); null models are Null4 and Null5. Dash line on top of the model bar is the level of the explained variance from the model without the environmental variables.

Table 4.3: Nestedness variance partitioning results showing the percentage each explanatory variable is contributing to the explained variance (R squared) in the observed values. SE is the sampling effort, ND is the nearest distance, AR is the area, HPS is the human population size (log transformed $[log(x+1)]$). Glm is the generalised linear model fitting; elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5.

		Variance partitioning				
Glm	R squared	Model	SE	ND	AR	HPS
Observed ~ IO + ND + AR + SE + HPS	0.79	1.15	2.04	3.42	8.70	84.69
Observed ~ IE + ND + AR + SE + HPS	0.86	30.46	2.61	4.27	6.92	55.74
Observed ~ IIO + ND + AR + SE + HPS	0.79	1.04	2.17	3.26	8.61	84.92
Observed ~ IIE + ND + AR + SE + HPS	0.85	30.06	2.66	4.25	6.98	56.04
Observed ~ Null4 + ND + AR + SE + HPS	0.92	30.65	6.36	6.34	6.00	50.65
Observed ~ Null5 + ND + AR + SE + HPS	0.83	22.04	2.72	4.72	6.99	63.54

4.3.2 Modularity

Here we start by stating the VIF results obtained from the explanatory variables and then we give the multiple regression results.

4.3.2.1 Variance Inflation Factor (VIF)

From the variance inflation factor (VIF) tests of all the explanatory variables which include the AIS and null model predictions, island area, nearest distance, age, maximum elevation, sampling effort and human population size, we dropped age and maximum elevation to maintain low VIF values less than 5 (Table 4.4).

Table 4.4: Modularity Variance Inflation Factor (VIF) showing the VIF values for each combination of the explanatory variables. ND is the nearest distance, SE is the sampling effort, AR is the area, HPS is the human population size (log transformed [$log(x+1)$]), ME is the maximum elevation, AG is the age. To obtain non collinearity among the variables, we make sure to achieve VIF values less than 5, dropping variables with high VIF values. (a) is the combination of all the environmental variables and the model predictions. (b) is the combination of the covariants without age. (c) is the combination of the covariants without age and maximum elevation. Elimination models are IE (type I, linear functional response) and IIE (type II, non-linear functional response); optimization models are IO (type I, linear functional response) and IIO (type II, non-linear functional response); null models are Null4 and Null5. R squared is the coefficient of determination of the Generalized Linear Model fittings.

Covariant	IO	IE	IIO	IIE	Null4	Null5
ND ~ AR + ME + HPS + AG + SE + model	3.29	4.9	2.76	4.81	6.99	5.88
AR ~ ND + ME + HPS + AG + SE + model	5.93	4.99	7.29	4.98	5.45	5.09
ME ~ ND + AR + HPS + AG + SE + model	4.16	4.91	3.86	4.86	7.24	5.57
HPS ~ ND + AR + ME + AG + SE + model	4.29	3.86	4.56	3.87	3.73	3.83
AG ~ ND + AR + ME + HPS + SE + model	8.14	9.99	5.22	9.7	15.89	11.8
SE ~ ND + AR + ME + HPS + AG + model	3.11	4.16	2.2	4.11	8.39	5.56
Model ~ ND + AR + ME + HPS + AG + SE	4.45	4.84	2.62	4.67	9.34	6.25
R squared of glm(observed ~ model + other variables)	0.97	0.97	0.97	0.97	0.97	0.97

Without age

Covariant	IO	IE	IIO	IIE	Null4	Null5
ND ~ AR + ME + HPS + SE + model	1.81	1.82	1.82	1.82	1.82	1.83
AR ~ ND + ME + HPS + SE + model	5.39	4.96	6.43	4.95	5.01	4.97
ME ~ ND + AR + HPS + SE + model	3.32	3.27	3.66	3.27	3.29	3.27
HPS ~ ND + AR + ME + SE + model	1.91	2.18	1.74	2.16	2.49	2.29
SE ~ ND + AR + ME + HPS + model	2.16	2.29	2.11	2.29	2.66	2.46
Model ~ ND + AR + ME + HPS + SE	1.66	1.48	1.53	1.47	1.79	1.61
R squared of glm(observed ~ model + other variables)	0.92	0.93	0.88	0.93	0.94	0.94

Without age and maximum elevation

Covariant	IO	IE	IIO	IIE	Null4	Null5
ND ~ AR + HPS + SE + model	1.79	1.8	1.79	1.8	1.81	1.82
AR ~ ND + HPS + SE + model	2.37	2.18	2.53	2.18	2.18	2.18
HPS ~ ND + AR + SE + model	1.83	2.07	1.63	2.05	2.33	2.16
SE ~ ND + AR + HPS + model	1.66	1.74	1.51	1.75	1.99	1.87
Model ~ ND + AR + HPS + SE	1.64	1.47	1.37	1.47	1.78	1.61
R squared of glm(observed ~ model + other variables)	0.89	0.91	0.85	0.91	0.92	0.92

4.3.2.2 Multiple regression

In general, the GLM fitting of the observed modularity values together with the remaining explanatory variables (model predictions, isolation, sampling effort, area and human population size) produced greater explained variance of the observed data. Various combinations explained significantly the observed level of modularity, from which we obtained the following: for the elimination algorithm, IE + other variables (R^2= 0.91, $p < 0.0001$), IIE + other variables

(R^2= 0.91, $p < 0.0001$); for optimization algorithm, IO + other variables (R^2= 0.89, $p < 0.0001$), IIO + other variables (R^2= 0.85, $p < 0.001$); for null models we obtained, Null4 + other variables (R^2= 0.92, $p < 0.0001$) and Null5 + other variables (R^2= 0.91, $p < 0.0001$) (Table 4.5). These suggest the important role the explanatory variables (environmental factors) are playing in explaining network structures.

Furthermore, we divided the total explained variance among the five explanatory variables. This allowed us to see the independent contribution of each of the covariant to the total explained variance (Fig. 4.2). In this partitioning, individual model contributions are IE = 32.9%, IIE = 33%, IO = 24.2%, IIO = 9.87%, Null4 = 39% and Null5 = 37% (Table 4.6). The sampling effort, nearest distances and the models were more important for modularity prediction. Area and human population size contributed a lesser percentage to the total explained variance in modularity (Table 4.6).

Table 4.5: Modularity multiple regression showing the Generalised linear model (Glm) fitting of the observed values with the model predictions, ND is the nearest distance, SE is the sampling effort, AR is the area, HPS is the human population size (log transformed [$log(x+1)$]). R squared is the coefficient of determination, AIC is the Akaike Information Criterion. Elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5.

Glm	R squared	AIC	p-values
Observed ~ IO + ND + AR + SE + HPS	0.8853	-39.2620	<0.0001
Observed ~ IE + ND + AR + SE + HPS	0.9081	-41.4800	<0.0001
Observed ~ IIO + ND + AR + SE + HPS	0.8453	-36.2710	0.0006
Observed ~ IIE + ND + AR + SE + HPS	0.9091	-41.5930	<0.0001
Observed ~ Null4 + ND + AR + SE + HPS	0.9235	-43.3180	<0.0001
Observed ~ Null5 + ND + AR + SE + HPS	0.9154	-42.3020	<0.0001

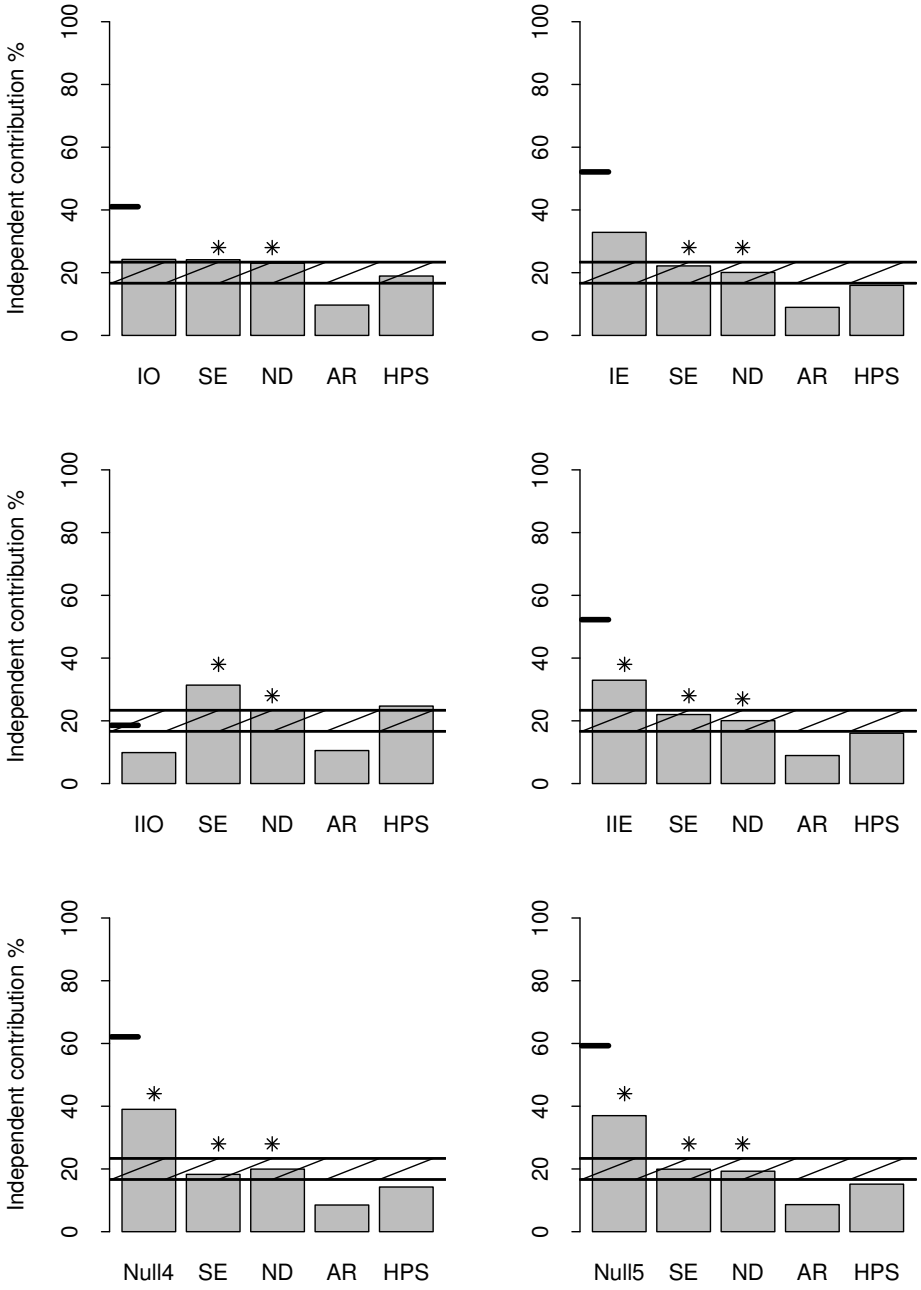


Figure 4.2: Modularity variance partitioning showing the independent contribution of each of the explanatory variables to the total variance explained for different model combinations. The environmental variables are SE - sampling effort, ND - nearest distance (isolation), AR - area and HPS - human population size (log transformed [log(x+1)]). The band strips shows the 95% confidence interval of the randomly divided variance explained. Bars above this band are important and asterisked bars are significant ($p < 0.1$). Elimination models are IE (linear functional response) and IIE (nonlinear functional response); optimization models are IO (linear functional response) and IIO (nonlinear functional response); null models are Null4 and Null5. Dash line on top of the model bar is the level of the explained variance from the model without the environmental variables.

Table 4.6: Modularity variance partitioning results showing the percentage each explanatory variable is contributing to the explained variance (R squared) in the observed values. SE is the sampling effort, ND is the nearest distance, AR is the area, HPS is the human population size (log transformed [$log(x+1)$]). Glm is the generalised linear model fitting; elimination models are IE (type I functional response) and IIE (type II functional response), optimization models are IO (type I functional response) and IIO (type II functional response), null models are Null4 and Null5.

		Variance partitioning				
Glm	R squared	Model	SE	ND	AR	HPS
Observed ~ IO + ND + AR + SE + HPS	0.89	24.23	24.13	23.04	9.67	18.93
Observed ~ IE + ND + AR + SE + HPS	0.91	32.86	22.16	20.08	8.95	15.96
Observed ~ IIO + ND + AR + SE + HPS	0.85	9.87	31.38	23.52	10.52	24.70
Observed ~ IIE + ND + AR + SE + HPS	0.91	32.95	22.01	20.06	8.94	16.05
Observed ~ Null4 + ND + AR + SE + HPS	0.92	39.01	18.27	19.96	8.51	14.24
Observed ~ Null5 + ND + AR + SE + HPS	0.92	36.99	19.93	19.29	8.63	15.16

4.4 Discussion

The role of island size, isolation, sampling effort and human disturbance are crucial in predicting network structures. The GLM regression we performed of the AIS models predictions together with these environmental variables tremendously increased the explained variation to more than 79% and more than 85% in the nestedness and modularity respectively. This suggests the importance of environmental in explaining structural emergence. Surprisingly, the null models with no special mechanism of interest, together with the environmental variables explained significantly more than 80% and more than 90% in nestedness and modularity respectively.

Geographic factors such as Island size and isolation play important role in determining the species richness on Islands (MacArthur and Wilson 1967; Roura-Pascual et al. 2016). From the theory of island biogeography, the bigger the island, the more the species it has. Although species richness may not directly affect nestedness and modularity, the effect is realised indirectly through the connectivity of the network. In essence, our results showed that the effect of island size and isolation on the network structure varies. Although island size had a lesser percentage contribution to nestedness and modularity, isolation was more important for modularity than for nestedness.

Human disturbances contributed more to the predicted nestedness of the network as evident from the variance partitioning of the GLM results. This suggests that human activities can lead to the introduction of exotic species which in turn lead to more realised interaction and then a more nested network. Our results is in consistent with Sebastian-Gonzalez et al. (2014). From the 55 seed

dispersal networks they analysed, networks with high human impact showed more nested and lesser modular structures. Sampling effort was more important for modularity than for nestedness predictions.

Chapter 5

Conclusions

Mutualistic interactions have played an important role in the maintenance of ecosystem functions and services. Therefore, structures that arise as a result of these interactions are crutial. In this chapter, we give a summary of our work in this thesis, highlighting the limitation in our study and further work that can be carried out.

5.1 Summary

Due to the importance of network structures, we have used Adaptive Interaction Switch (AIS) models to predict the emergence of network structures from the data collected from 10 Galápagos island networks. More importantly, we predicted for nestedness and modularity network structures. Owing to the ecological and evolutionary processes incorporated in the Lotka-Voltera models, an appreciable level of variation in the observed nestedness and modularity was explained.

Adding the environmental variables to the AIS model predictions, our Generalized Linear Model (GLM) results showed a high level of explained variation in the observed structure. This showed the importance of the environmental variables as predictors of network structures, which included island area, isolation, sampling effort and human population size. Hence the eco-evolutionary process and environmental variables are potential key drivers of structural emergence in mutalistic networks. These results, on a crucial note, are fundamental to ecological network studies, providing more understanding to environmental and anthropogenic disturbances to the ecosystem.

5.2 Limitation and further work

As a limitation, quantitative data was converted to binary networks. Although we have shown emergence of network structures using a binary interaction matrix depicting presence of an interaction (1) or not (0) between the plants and pollinators, the results were robust. Hence, a more integrative analysis can be done using weighted networks, taking into account frequency of interactions. Investigating the emergence of structures connected by propagules, moving in-between local communities can be studied.

Appendices

Appendix A

R codes

Below we give the R codes for the Adaptive Interaction Switch models.

```
## PARAMETER DEFINITIONS

# M    - Number of plants.
# N    - Number of animals.
# C    - Number of interactions.
# A    - M x N binary interaction matrix.
# rP   - M x 1 matrix for plants intrinsic growth rates.
# rA   - N x 1 matrix for animals intrinsic growth rates.
# PP   - M x M matrix for plants intra (diagonal elements)
#        and inter specific competition strengths.
# AA   - N x N matrix for animals intra (diagonal elements)
#        and inter specific competition strengths.
# PA   - M x N matrix for strength of mutualistic benefits
#        of plants from animals.
# AP   - N x M matrix for strength of mutualistic benefits
#        of animals from plants.
# h    - Handling time, h=0 and h=0.1 for type I and type II
#        functional response respectively.
# XP0  - M x 1 matrix for plants initial population sizes.
# XA0  - N x 1 matrix for animals initial population sizes.
# T    - Number of switching events.
# v    - M x N matrix for preference values.


## R PACKAGES NEEDED.
library(bipartite)
library(deSolve)
```

```
## LOTKA-VOLTERA EQUATIONS.

lotka <- function(tt,y,parameters){
  with(as.list(c(y,parameters)),{
    Xx <- y
    M <- dim(A)[1]
    N <- dim(A)[2]
    XP <- Xx[1:M]
    XA <- Xx[(M+1):(M+N)]
    dXP <- rP - PP %*% XP + ((PA * A * v) %*% XA / (1 + (h*A*v)%*%XA))
    dXA <- rA - AA %*% XA + ((AP * t(A*v)) %*% XP / (1 + (h*t(A*v))%*%XP))
    dX <- Xx * rbind(dXP,dXA)
    return(list(dX))
  })
}

## NESTEDNESS FUNCTION
# From bipartite package

NODF <- function(dt){
  A <- nested(dt, method = "NODF2", rescale=FALSE, normalised=TRUE)/100
  return (A)
}


## MODULARITY FUNCTION
# From Netcarto software, implemented in R

modularity <- function(A){

  A_index <- which(A != 0, arr.ind = TRUE)
  write.table(A_index, file="A_index.dat", row.names = FALSE, col.names = FALSE)
  system("wine netcarto_cl A_index.dat 1 5.0 0.1 0.95 0")
  x <- scan("modules.dat", what = list(Modularity = ""))
  mod_A <- as.numeric(unlist(x)[length(unlist(x))])
  return(mod_A)
}




## ELIMINATION ALGORITHM
```

```
# From  Zhang et al. 2011 An interaction switch predicts the nested
# architecture of mutualistic networks. Ecol. Lett. 14, 797-803.

elim <- function(A,rP,rA,PP,AA,PA,AP,h,XP0,XA0,T,v){

  M = dim(A)[1]
  N = dim(A)[2]

  X = matrix(0,M+N,T)       # Species population sizes
  X[,1] = rbind(XP0,XA0)

  NEST <- matrix(0,1,T)     # Nestedness
  NEST[1] <- NODF(A)

  MOD <- matrix(0,1,T)      # Modularity
  MOD[1] <- modularity(A)

  for (t in 2:T){
    yini = X[,t-1]
    times = seq(0,1,0.01)
    parameters = list(A=A,rP=rP,rA=rA,PP=PP,AA=AA,PA=PA,AP=AP,h=h,v=v)
    out <- rk4(y = yini, times = times, func=lotka, parms=parameters)
    X[,t] <- as.matrix(out[length(times),-1])

    mn <- sample(M+N,1)
    if (mn <= M){
      m <- mn
      Benefit <- t(PA[m,]*A[m,]*v[m,]) * X[(M+1):(M+N),t]
      B_min <- Benefit[Benefit > 0]
      n1 <- which(Benefit==min(B_min))

      if (sum(A[,n1])>1){
        Am <- A[m,]
        rm0 <- which(Am==0)

        if (length(rm0) != 0){
          n0 <- rm0[sample(length(rm0),1)]
          A[m,c(n1,n0)] <- A[m,c(n0,n1)]
          v[m,c(n1,n0)] <- v[m,c(n0,n1)]

          NEST[t] <- NODF(A)
          MOD[t] <- modularity(A)

        } else{
```

```
          NEST[t] <- NEST[t-1]
          MOD[t] <- MOD[t-1]
        }
      } else{
        NEST[t] <- NEST[t-1]
        MOD[t] <- MOD[t-1]
      }
    } else{
      n <- mn - M
      Benefit <- t(t(AP)[,n]*A[,n]*v[,n]) * X[1:M,t]
      B_min <- Benefit[Benefit > 0]
      m1 <- which(Benefit==min(B_min))

      if (sum(A[m1,])>1){
        An <- A[,n]
        rn0 <- which(An==0)

        if (length(rn0) != 0){
          m0 <- rn0[sample(length(rn0),1)]
          A[c(m1,m0),n] <- A[c(m0,m1),n]
          v[c(m1,m0),n] <- v[c(m0,m1),n]

          NEST[t] <- NODF(A)
          MOD[t] <- modularity(A)

        } else{
          NEST[t] <- NEST[t-1]
          MOD[t] <- MOD[t-1]
        }

      } else{
        NEST[t] <- NEST[t-1]
        MOD[t] <- MOD[t-1]
      }
    }
  }
  return(list("X"=X, "Nestedness" = NEST, "Modularity" = MOD, "A"=A))
}


## OPTIMIZATION ALGORITHM

# From Suweis et al. 2013 Emergence of structural and dynamical properties
# of ecological mutualistic networks. Nature 500, 449-452.
```

```
opt <- function(A,rP,rA,PP,AA,PA,AP,h,XP0,XA0,T,v){

  M = dim(A)[1]
  N = dim(A)[2]

  X = matrix(0,M+N,T)       # Species population sizes
  X[,1] = rbind(XP0,XA0)

  NEST <- matrix(0,1,T)     # Nestedness
  NEST[1] <- NODF(A)

  MOD <- matrix(0,1,T)      # Modularity
  MOD[1] <- modularity(A)

  for(t in 2:T){
    mn <- sample(M+N,1)

    if (mn <= M){
      m <- mn
      Am <- A[m,]
      rm1 <- which(Am==1)
      n1 <- rm1[sample(length(rm1),1)]
      rm0 <- 1:N ; rm0 <- rm0[-n1]
      n0 <- rm0[sample(length(rm0),1)]
      A[m,c(n1,n0)] <- A[m,c(n0,n1)]
      PA[m,c(n1,n0)] <- PA[m,c(n0,n1)]
      AP[c(n1,n0),m] <- AP[c(n0,n1),m]
      v[m,c(n1,n0)] <- v[m,c(n0,n1)]

      yini <- X[,t-1]
      times <- seq(0,1,0.01)
      parameters = list(A=A,rP=rP,rA=rA,PP=PP,AA=AA,PA=PA,AP=AP,h=h,v=v)
      out <- rk4(y = yini, times = times, func=lotka, parms=parameters)
      X[,t] <- as.matrix(out[length(times),-1])

      if (X[mn,t] >= X[mn,(t-1)] && sum(A[,n1])>0 && sum(A[,n0])>0){
        NEST[t] <- NODF(A)
        MOD[t] <- modularity(A)

      } else{
        A[m,c(n1,n0)] <- A[m,c(n0,n1)]
        PA[m,c(n1,n0)] <- PA[m,c(n0,n1)]
        AP[c(n1,n0),m] <- AP[c(n0,n1),m]
```

```
      v[m,c(n1,n0)] <- v[m,c(n0,n1)]
      X[,t] <- X[,t-1]
      NEST[t] <- NEST[t-1]
      MOD[t] <- MOD[t-1]
    }

  } else{

    n <- mn - M
    An <- A[,n]
    rn1 <- which(An==1)
    m1 <- rn1[sample(length(rn1),1)]
    rn0 <- 1:M ; rn0 <- rn0[-m1]
    m0 <- rn0[sample(length(rn0),1)]
    A[c(m1,m0),n] <- A[c(m0,m1),n]
    PA[c(m1,m0),n] <- PA[c(m0,m1),n]
    AP[n,c(m1,m0)] <- AP[n,c(m0,m1)]
    v[c(m1,m0),n] <- v[c(m0,m1),n]

    yini <- X[,t-1]
    times <- seq(0,1,0.01)
    parameters = list(A=A,rP=rP,rA=rA,PP=PP,AA=AA,PA=PA,AP=AP,h=h,v=v)
    out <- rk4(y = yini, times = times, func=lotka, parms=parameters)
    X[,t] <- as.matrix(out[length(times),-1])


    if (X[mn,t] >= X[mn,(t-1)] && sum(A[m1,])>0 && sum(A[m0,])>0){
      NEST[t] <- NODF(A)
      MOD[t] <- modularity(A)

    }else {

      A[c(m1,m0),n] <- A[c(m0,m1),n]
      PA[c(m1,m0),n] <- PA[c(m0,m1),n]
      AP[n,c(m1,m0)] <- AP[n,c(m0,m1)]
      v[c(m1,m0),n] <- v[c(m0,m1),n]
      X[,t] <- X[,t-1]
      NEST[t] <- NEST[t-1]
      MOD[t] <- MOD[t-1]
    }
  }
}

return(list("X"=X, "Nestedness" = NEST, "Modularity" = MOD, "A"=A))
```

```
}

#############################################################################


## A DEMONSTRATION WITH ESPANOLA ISLAND DATA

M = 11
N = 16
C = 31
S = M * N     # Network size.

# Randomly generate binary interaction matrix.
AR = matrix(1,M,N)
while (S>C){
  m <- sample(M,1)
  n <- sample(N,1)
  if (AR[m,n]>0 && sum(AR[m,])>1 && sum(AR[,n])>1){
    AR[m,n] <- 0
    S <- S-1
  }
}


#### Assign values to parameters.

T <- 5000

sigma <- 0.01 # Standard deviation value for the normal distributions.

# Generate competition matrix
PP <- abs(matrix(rnorm(M*M,0,sigma), nr=M))
PP <- PP - diag(diag(PP)) + diag(diag(matrix(1,M,M)))
AA <- abs(matrix(rnorm(N*N,0,sigma), nr=N))
AA <- AA - diag(diag(AA)) + diag(diag(matrix(1,N,N)))

# Generate mutualistic benefit matrix
PA <- abs(matrix(rnorm(M*N,0,sigma), nr=M))
AP <- abs(matrix(rnorm(M*N,0,sigma), nr=N))

# Generate initial population sizes
XP0 <- matrix(runif(M,0,1), nr= M)
XA0 <- matrix(runif(N,0,1), nr= N)
```

```
# Generate intrinsic growth sizes
rP <- matrix(rlnorm(M,0,0.1))
rA <- matrix(rlnorm(N,0,0.1))

# Generate preference matrix
v <- matrix(runif(M*N),nr=M)



#### MODEL WITH TYPE I FUNCTIONAL RESPONSE.
h <- 0
A <- AR

#* ELIMINATION ALGORITHM
IE <- elim(A,rP,rA,PP,AA,PA,AP,h,XP0,XA0,T,v);
IE_nest <- matrix(unlist(IE[2]),1,T)
IE_mod <- matrix(unlist(IE[3]),1,T)

#* OPTIMIZATION ALGORITHM
IO <- opt(A,rP,rA,PP,AA,PA,AP,h,XP0,XA0,T,v)
IO_nest <- matrix(unlist(IO[2]),1,T)
IO_mod <- matrix(unlist(IO[3]),1,T)


#### MODEL WITH TYPE II FUNCTIONAL RESPONSE.
h <- 0.1
A <- AR

#* ELIMINATION ALGORITHM
IIE <- elim(A,rP,rA,PP,AA,PA,AP,h,XP0,XA0,T,v);
IIE_nest <- matrix(unlist(IIE[2]),1,T)
IIE_mod <- matrix(unlist(IIE[3]),1,T)

#* OPTIMIZATION ALGORITHM
IIO <- opt(A,rP,rA,PP,AA,PA,AP,h,XP0,XA0,T,v)
IIO_nest <- matrix(unlist(IIO[2]),1,T)
IIO_mod <- matrix(unlist(IIO[3]),1,T)


## PLOTS

# NESTEDNESS PREDICTONS

plot(1:T,seq(0,0.8,len=T), type = "n", xlab = "Time",ylab = "Nestedness")
```

```
lines(1:T, IO_nest, col = "blue",lwd=2)
lines(1:T, IE_nest, col = "green",lwd=2)
lines(1:T, IIO_nest, col = "purple",lwd=2)
lines(1:T, IIE_nest, col = "yellow",lwd=2)
lines(1:T, rep(0.2076,T), col= "red",lwd=2,lty=2)
legend('topleft', c("IO","IE","IIO","IIE","Observed"), lty = c(1,1,1,1,2),
       lwd = c(2,2,2,2), col=c("blue","green","purple","yellow","red"),
       cex = 0.8, bty = "n")


# MODULARITY PREDICTONS

plot(1:T,seq(0,0.8,len=T), type = "n", xlab = "Time",ylab = "Modularity")
lines(1:T, IO_mod, col = "blue",lwd=2)
lines(1:T, IE_mod, col = "green",lwd=2)
lines(1:T, IIO_mod, col = "purple",lwd=2)
lines(1:T, IIE_mod, col = "yellow",lwd=2)
lines(1:T, rep(0.3592,T), col= "red",lwd=2,lty=2)
legend('topleft', c("IO","IE","IIO","IIE","Observed"), lty = c(1,1,1,1,2),
       lwd = c(2,2,2,2), col=c("blue","green","purple","yellow","red"),
       cex = 0.8, bty = "n")
```

List of References

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, vol. 74, no. 1, p. 47.

Albert, R., Jeong, H. and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, vol. 406, no. 6794, pp. 378–382.

Allesina, S. and Tang, S. (2012). Stability criteria for complex ecosystems. *Nature*, vol. 483, no. 7388, pp. 205–208.

Almeida-Neto, M., Guimaraes, P., Guimarães, P.R., Loyola, R.D. and Ulrich, W. (2008). A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, vol. 117, no. 8, pp. 1227–1239.

Alonso, D., Etienne, R.S. and McKane, A.J. (2006). The merits of neutral theory. *Trends in Ecology & Evolution*, vol. 21, no. 8, pp. 451–457.

Amaral, L.A.N., Scala, A., Barthelemy, M. and Stanley, H.E. (2000). Classes of small-world networks. *Proceedings of the national academy of sciences*, vol. 97, no. 21, pp. 11149–11152.

Atmar, W. and Patterson, B.D. (1993). The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia*, vol. 96, no. 3, pp. 373–382.

Bar-Yam, S. (2016*a*). Plant herbivory. http://necsi.edu/projects/evolution/darwin/beagle/Galapagos/darwin_beagle_galapagos.html. Accessed: 2016-06-22.

Bar-Yam, S. (2016*b*). Predator prey. http://necsi.edu/projects/evolution/co-evolution/pred-prey/co-evolution_predator.html. Accessed: 2016-06-22.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *science*, vol. 286, no. 5439, pp. 509–512.

Bascompte, J. and Jordano, P. (2006). The structure of plant-animal mutualistic networks. *Ecological networks: linking structure to dynamics in food webs. Oxford University Press, Oxford, UK*, pp. 143–159.

Bascompte, J., Jordano, P., Melián, C.J. and Olesen, J.M. (2003). The nested assembly of plant–animal mutualistic networks. *Proceedings of the National Academy of Sciences*, vol. 100, no. 16, pp. 9383–9387.

Bascompte, J., Jordano, P. and Olesen, J.M. (2006). Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science*, vol. 312, no. 5772, pp. 431–433.

Basilio, A.M., Medan, D., Torretta, J.P. and Bartoloni, N.J. (2006). A year-long plant-pollinator network. *Austral Ecology*, vol. 31, no. 8, pp. 975–983.

Bastolla, U., Fortuna, M.A., Pascual-García, A., Ferrera, A., Luque, B. and Bascompte, J. (2009). The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, vol. 458, no. 7241, pp. 1018–1020.

Batra, S.W.T. (2016). Solitary bees. http://www.pollinatorparadise.com/solitary_bees/Solitar.htm. Accessed: 2016-06-22.

Bollobás, B. (1998). Random graphs. In: *Modern Graph Theory*, pp. 215–252. Springer.

Burgos, E., Ceva, H., Perazzo, R.P., Devoto, M., Medan, D., Zimmermann, M. and Delbue, A.M. (2007). Why nestedness in mutualistic networks? *Journal of theoretical biology*, vol. 249, no. 2, pp. 307–313.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, pp. 343–366.

Cole, M. (2016). Marmalade hoverfly. http://www.mattcolephotography.co.uk/Galleries/insects/Hoverflies/slides/Marmalade%20Hoverfly%206.html. Accessed: 2016-06-22.

Conservancy, G. (2016). Galapagos conservancy. http://www.galapagos.org/about_galapagos/about-galapagos/the-islands. Accessed: 2016-05-21.

Dicks, L., Corbet, S. and Pywell, R. (2002). Compartmentalization in plant–insect flower visitor webs. *Journal of Animal Ecology*, vol. 71, no. 1, pp. 32–43.

Donohue, I., Petchey, O.L., Montoya, J.M., Jackson, A.L., McNally, L., Viana, M., Healy, K., Lurgi, M., O'Connor, N.E. and Emmerson, M.C. (2013). On the dimensionality of ecological stability. *Ecology letters*, vol. 16, no. 4, pp. 421–429.

Dunne, J.A., Williams, R.J. and Martinez, N.D. (2002). Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Sciences*, vol. 99, no. 20, pp. 12917–12922.

Ehrlich, P.R. and Raven, P.H. (1964). Butterflies and plants: a study in coevolution. *Evolution*, pp. 586–608.

Erdös, P. and Rényi, A. (1959). On random graphs, i. *Publicationes Mathematicae (Debrecen)*, vol. 6, pp. 290–297.

Erdös, P. and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, no. 17-61, p. 43.

Erdös, P. and Rényi, A. (1961). On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, vol. 12, no. 1-2, pp. 261–267.

Feng, W. and Takemoto, K. (2014). Heterogeneity in ecological mutualistic networks dominantly determines community stability. *Scientific reports*, vol. 4.

Fortuna, M.A. and Bascompte, J. (2006). Habitat loss and the structure of plant–animal mutualistic networks. *Ecology Letters*, vol. 9, no. 3, pp. 281–286.

Fortuna, M.A., Stouffer, D.B., Olesen, J.M., Jordano, P., Mouillot, D., Krasnov, B.R., Poulin, R. and Bascompte, J. (2010). Nestedness versus modularity in ecological networks: two sides of the same coin? *Journal of Animal Ecology*, vol. 79, no. 4, pp. 811–817.

Fossette, S., Gleiss, A.C., Casey, J.P., Lewis, A.R. and Hays, G.C. (2011). Does prey size matter? novel observations of feeding in the leatherback turtle (dermochelys coriacea) allow a test of predator–prey size relationships. *Biology Letters*, p. rsbl20110965.

Gilljam, D., Curtsdotter, A. and Ebenman, B. (2015). Adaptive rewiring aggravates the effects of species loss in ecosystems. *Nature Communications*, vol. 6.

Gotelli, N.J.G. *et al.* (1996). *Null models in ecology*. 574.501519 G6.

Guimarães, P.R., Rico-Gray, V., Dos Reis, S.F. and Thompson, J.N. (2006). Asymmetries in specialization in ant–plant mutualistic networks. *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 273, no. 1597, pp. 2041–2047.

Guimarães, P.R., Sazima, C., Dos Reis, S.F. and Sazima, I. (2007). The nested structure of marine cleaning symbiosis: is it like flowers and bees? *Biology Letters*, vol. 3, no. 1, pp. 51–54.

Guimera, R. and Amaral, L.A.N. (2005). Functional cartography of complex metabolic networks. *Nature*, vol. 433, no. 7028, pp. 895–900.

Herrera, H. and Roque-Álbelo, L. (2012). Cdf checklist of galapagos terrestrial invertebrates?fcd lista de especies de invertebrados terrestres de galápagos. *Charles Darwin Foundation Galapagos Species Checklist. Charles Darwin Foundation, Puerto Ayora, Galápagos: http://www. darwinfoundation. org/datazone/checklists/terrestrial-invertebrates/Last updated*, vol. 29.

Holland, J.N., DeAngelis, D.L. and Bronstein, J.L. (2002). Population dynamics and mutualism: functional responses of benefits and costs. *The American Naturalist*, vol. 159, no. 3, pp. 231–244.

Holland, J.N., Okuyama, T. and DeAngelis, D.L. (2006). Comment on "asymmetric coevolutionary networks facilitate biodiversity maintenance". *Science*, vol. 313, no. 5795, pp. 1887b–1887b.

Holling, C.S. (1959). Some characteristics of simple types of predation and parasitism. *The Canadian Entomologist*, vol. 91, no. 07, pp. 385–398.

Ives, A.R. and Carpenter, S.R. (2007). Stability and diversity of ecosystems. *science*, vol. 317, no. 5834, pp. 58–62.

James, A., Pitchford, J.W. and Plank, M.J. (2012). Disentangling nestedness from models of ecological complexity. *Nature*, vol. 487, no. 7406, pp. 227–230.

Jaramillo, P., Guézou, A., Mauchamp, A. and Tye, A. (2012). Cdf checklist of galapagos flowering plants. *Charles Darwin Foundation Galapagos Species Checklist*.

Jordano, P. (1987). Patterns of mutualistic interactions in pollination and seed dispersal: connectance, dependence asymmetries, and coevolution. *American naturalist*, pp. 657–677.

Jordano, P., Bascompte, J. and Olesen, J.M. (2003). Invariant properties in coevolutionary networks of plant–animal interactions. *Ecology Letters*, vol. 6, no. 1, pp. 69–81.

Jordano, P. *et al.* (2000). Fruits and frugivory. *Seeds: the ecology of regeneration in plant communities*, vol. 2, pp. 125–166.

Kaiser-Bunbury, C.N., Muff, S., Memmott, J., Müller, C.B. and Caflisch, A. (2010). The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour. *Ecology Letters*, vol. 13, no. 4, pp. 442–452.

Karoński, M. (1982). A review of random graphs. *Journal of Graph Theory*, vol. 6, no. 4, pp. 349–389.

Kondoh, M. (2003). Foraging adaptation and the relationship between food-web complexity and stability. *Science*, vol. 299, no. 5611, pp. 1388–1391.

Latora, V. and Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical review letters*, vol. 87, no. 19, p. 198701.

Lewinsohn, T.M., Inácio Prado, P., Jordano, P., Bascompte, J. and M Olesen, J. (2006). Structure in plant–animal interaction assemblages. *Oikos*, vol. 113, no. 1, pp. 174–184.

MacArthur, R. and Wilson, E. (1967). The theory of island biography. *Princeton, New Jersey*.

Manaster, J. (2016). Interspecies competition. http://blogs.scientificamerican.com/psi-vid/mammal-march-madness/. Accessed: 2016-06-22.

May, R.M. (1972). Will a large complex system be stable? *Nature*, vol. 238, pp. 413–414.

May, R.M. (1973). *Stability and complexity in model ecosystems*, vol. 6. Princeton University Press.

May, R.M. (1981). Models of two interacting populations. *Theoretical ecology*, pp. 78–104.

McCann, K.S. (2000). The diversity–stability debate. *Nature*, vol. 405, no. 6783, pp. 228–233.

Memmott, J. (1999). The structure of a plant-pollinator food web. *Ecology Letters*, vol. 2, no. 5, pp. 276–280.

Milgram, S. (1967). The small world problem. *Psychology Today*, vol. 2, no. 1, pp. 60–67.

Minoarivelo, H.O., Hui, C., Terblanche, J., Kosakovsky Pond, S. and Scheffler, K. (2014). Detecting phylogenetic signal in mutualistic interaction networks using a markov process model. *Oikos*, vol. 123, no. 10, pp. 1250–1260.

Mohammed, S. (2016). Intraspecies competition. https://za.pinterest.com/pin/408772103648134301/. Accessed: 2016-06-22.

Montoya, J.M., Pimm, S.L. and Solé, R.V. (2006). Ecological networks and their fragility. *Nature*, vol. 442, no. 7100, pp. 259–264.

Moore, P.D. (2001). Ecology: The guts of seed dispersal. *Nature*, vol. 414, no. 6862, pp. 406–407.

Newman, M., Barabasi, A. and Watts, D.J. (2011). *The structure and dynamics of networks*. Princeton University Press.

Newman, M.E. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, vol. 69, no. 2, p. 026113.

Nogales, M., Heleno, R., Rumeu, B., González-Castro, A., Traveset, A., Vargas, P. and Olesen, J. (2015). Seed-dispersal networks on the canaries and the galápagos archipelagos: interaction modules as biogeographical entities. *Global Ecology and Biogeography*.

Nuismer, S.L., Jordano, P. and Bascompte, J. (2013). Coevolution and the architecture of mutualistic networks. *Evolution*, vol. 67, no. 2, pp. 338–354.

Nuwagaba, S., Zhang, F. and Hui, C. (2015). A hybrid behavioural rule of adaptation and drift explains the emergent architecture of antagonistic networks. In: *Proc. R. Soc. B*, vol. 282, p. 20150320. The Royal Society.

Okuyama, T. and Holland, J.N. (2008). Network structural properties mediate the stability of mutualistic communities. *Ecology Letters*, vol. 11, no. 3, pp. 208–216.

Olesen, J.M., Bascompte, J., Dupont, Y.L. and Jordano, P. (2007). The modularity of pollination networks. *Proceedings of the National Academy of Sciences*, vol. 104, no. 50, pp. 19891–19896.

Ollerton, J., McCollin, D., Fautin, D.G. and Allen, G.R. (2007). Finding nemo: nestedness engendered by mutualistic organization in anemonefish and their hosts. *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 274, no. 1609, pp. 591–598.

Paine, R.T. (1963). Trophic relationships of 8 sympatric predatory gastropods. *Ecology*, vol. 44, no. 1, pp. 63–73.

Paine, R.T. (1980). Food webs: linkage, interaction strength and community infrastructure. *Journal of Animal Ecology*, vol. 49, no. 3, pp. 667–685.

Pascual, M. and Dunne, J.A. (2006). *Ecological networks: linking structure to dynamics in food webs*. Oxford University Press.

Petanidou, T., Kallimanis, A.S., Tzanopoulos, J., Sgardelis, S.P. and Pantis, J.D. (2008). Long-term observation of a pollination network: fluctuation in species and interactions, relative invariance of network structure and implications for estimates of specialization. *Ecology Letters*, vol. 11, no. 6, pp. 564–575.

Pimm, S.L. (1979). Complexity and stability: another look at macarthur's original hypothesis. *Oikos*, pp. 351–357.

Pimm, S.L. (1984). The complexity and stability of ecosystems. *Nature*, vol. 307, no. 5949, pp. 321–326.

Poulakakis, N., Russello, M., Geist, D. and Caccone, A. (2012). Unravelling the peculiarities of island life: vicariance, dispersal and the diversification of the extinct and extant giant galápagos tortoises. *Molecular ecology*, vol. 21, no. 1, pp. 160–173.

Proctor, M., Yeo, P., Lack, A. *et al.* (1996). *The natural history of pollination*. HarperCollins Publishers.

Rapoport, A. (1957). Contribution to the theory of random and biased nets. *The Bulletin of Mathematical Biophysics*, vol. 19, no. 4, pp. 257–277.

Rezende, E.L., Lavabre, J.E., Guimarães, P.R., Jordano, P. and Bascompte, J. (2007). Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature*, vol. 448, no. 7156, pp. 925–928.

Rivera-Hutinel, A., Bustamante, R., Marín, V. and Medel, R. (2012). Effects of sampling completeness on the structure of plant–pollinator networks. *Ecology*, vol. 93, no. 7, pp. 1593–1603.

Rominger, A., Goodman, K., Lim, J., Armstrong, E., Becking, L., Bennett, G., Brewer, M., Cotoras, D., Ewing, C., Harte, J. *et al.* (2015). Community assembly on isolated islands: macroecology meets evolution. *Global Ecology and Biogeography*.

Rosenzweig, M.L. and MacArthur, R.H. (1963). Graphical representation and stability conditions of predator-prey interactions. *American Naturalist*, pp. 209–223.

Roura-Pascual, N., Sanders, N.J. and Hui, C. (2016). The distribution and diversity of insular ants: do exotic species play by different rules? *Global Ecology and Biogeography*, vol. 25, no. 6, pp. 642–654.

Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A. *et al.* (2000). Global biodiversity scenarios for the year 2100. *Science*, vol. 287, no. 5459, pp. 1770–1774.

Sebastián-González, E., Dalsgaard, B., Sandel, B. and Guimarães, P.R. (2014). Macroecological trends in nestedness and modularity of seed-dispersal networks: human impact matters. *Global Ecology and Biogeography*, vol. 24, no. 3, pp. 293–303.

Severns, J. (2016). European honey bee. https://commons.wikimedia.org/wiki/File:European_honey_bee_extracts_nectar.jpg. Accessed: 2016-06-22.

Sole, R.V. and Montoya, M. (2001). Complexity and fragility in ecological networks. *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 268, no. 1480, pp. 2039–2045.

Solomonoff, R. and Rapoport, A. (1951). Connectivity of random nets. *The Bulletin of Mathematical Biophysics*, vol. 13, no. 2, pp. 107–117.

Staniczenko, P., Lewis, O.T., Jones, N.S. and Reed-Tsochas, F. (2010). Structural dynamics and robustness of food webs. *Ecology Letters*, vol. 13, no. 7, pp. 891–899.

Stephens, D.W. and Krebs, J.R. (1986). *Foraging Theory*. Princeton University Press.

Suweis, S., Simini, F., Banavar, J.R. and Maritan, A. (2013). Emergence of structural and dynamical properties of ecological mutualistic networks. *Nature*, vol. 500, no. 7463, pp. 449–452.

Thébault, E. and Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, vol. 329, no. 5993, pp. 853–856.

Thompson, J.N. (1994). *The coevolutionary process*. University of Chicago Press.

Thompson, J.N. (2005). *The geographic mosaic of coevolution*. University of Chicago Press.

Tilman, D. (1982). *Resource competition and community structure*. 17. Princeton university press.

Traveset, A., Heleno, R., Chamorro, S., Vargas, P., McMullen, C.K., Castro-Urgal, R., Nogales, M., Herrera, H.W. and Olesen, J.M. (2013). Invaders of pollination networks in the galápagos islands: emergence of novel communities. *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 280, no. 1758, p. 20123040.

Tribune, R. (2016). Bumblebee. http://www.regaltribune.com/pollen-nectar-flight-bumblebee/22231/. Accessed: 2016-06-22.

Trøjelsgaard, K., Báez, M., Espadaler, X., Nogales, M., Oromí, P., Roche, F.L. and Olesen, J.M. (2013). Island biogeography of mutualistic interaction networks. *Journal of Biogeography*, vol. 40, no. 11, pp. 2020–2031.

Tutorvista (2016). Bird pollination. http://biology.tutorvista.com/plant-kingdom/pollination.html. Accessed: 2016-06-22.

Tylianakis, J.M., Didham, R.K., Bascompte, J. and Wardle, D.A. (2008). Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, vol. 11, no. 12, pp. 1351–1363.

Tylianakis, J.M., Tscharntke, T. and Lewis, O.T. (2007). Habitat modification alters the structure of tropical host–parasitoid food webs. *Nature*, vol. 445, no. 7124, pp. 202–205.

Valdovinos, F.S., Ramos-Jiliberto, R., Garay-Narváez, L., Urbani, P. and Dunne, J.A. (2010). Consequences of adaptive behaviour for the structure and dynamics of food webs. *Ecology Letters*, vol. 13, no. 12, pp. 1546–1559.

Vázquez, D.P. and Aizen, M.A. (2004). Asymmetric specialization: a pervasive feature of plant–pollinator interactions. *Ecology*, vol. 85, no. 5, pp. 1251–1257.

Vázquez, D.P., Blüthgen, N., Cagnolo, L. and Chacoff, N.P. (2009). Uniting pattern and process in plant–animal mutualistic networks: a review. *Annals of Botany*, vol. 103, no. 9, pp. 1445–1457.

Vázquez, D.P., Melián, C.J., Williams, N.M., Blüthgen, N., Krasnov, B.R. and Poulin, R. (2007). Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, vol. 116, no. 7, pp. 1120–1127.

Vázquez, D.P., Morris, W.F. and Jordano, P. (2005*a*). Interaction frequency as a surrogate for the total effect of animal mutualists on plants. *Ecology Letters*, vol. 8, no. 10, pp. 1088–1094.

Vázquez, D.P., Poulin, R., Krasnov, B.R. and Shenbrot, G.I. (2005*b*). Species abundance and the distribution of specialization in host–parasite interaction networks. *Journal of Animal Ecology*, vol. 74, no. 5, pp. 946–955.

Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature*, vol. 393, no. 6684, pp. 440–442.

Whittall, J.B. and Hodges, S.A. (2007). Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature*, vol. 447, no. 7145, pp. 706–709.

Zhang, F. and Hui, C. (2014). Recent experience-driven behaviour optimizes foraging. *Animal Behaviour*, vol. 88, pp. 13–19.

Zhang, F., Hui, C. and Terblanche, J.S. (2011). An interaction switch predicts the nested architecture of mutualistic networks. *Ecology Letters*, vol. 14, no. 8, pp. 797–803.